GENERATION AND ANALYSES OF EXPRESSED SEQUENCE TAGS FROM *SARGASSUM POLYCYSTUM* C. AGARDH (FUCALES, OCHROPHYTA)

SIM MEI CHEA

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2016

GENERATION AND ANALYSES OF EXPRESSED SEQUENCE TAGS FROM *SARGASSUM POLYCYSTUM* C. AGARDH (FUCALES, OCHROPHYTA)

SIM MEI CHEA

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2016

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: SIM MEI CHEA

Matric No: SHC 070039

Name of Degree: DOCTOR OF PHILOSOPHY

Title of Project Paper/Research Report/Dissertation/Thesis: (―Generation and Analyses of Expressed Sequence Tags from *Sargassum polycystum* C. Agardh (Fucales, Ochrophyta)")

Field of Study: Algal Biotechnology

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya (―UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

ABSTRACT

Sargassum polycystum C. Agardh is an economically and ecologically important brown seaweed, being one of the main raw materials for the production of alginate in Asian countries with many applications in the food, feed, pharmaceutical and medical industries. However, its entire genome has not been sequenced, and limited resources are available in the GenBank for understanding the molecular mechanisms underlying alginate biosynthesis. In this study, 2577 high-quality expressed sequence tags (ESTs) were generated from the cDNA library derived from pooled RNA isolated from *S. polycystum* samples. The ESTs were assembled with stringent parameters, resulting in 295 contigs and 1429 singletons, and giving a total of 1724 unigenes. Of all proteins putatively coded by these unigenes, 46.3 % have significant hits to known proteins in the Swiss-Prot database, and 66.2 % showed significant similarity to sequences in the NCBI non-redundant protein database. The unigenes were functionally characterized by Gene Ontology (GO) annotation, in which ~ 40.3 % were classified into one or more GO categories. We identified two unigenes encoding a potential alginate-related enzyme, mannuronan C5-epimerase (MC5E) and two unigenes encoding sulfated fucan-related enzyme, GDP mannose 4,6 dehyratase (GM46D), and GDP-4-keto-6-D-mannose-epimerase-4-reductase (GFS), respectively. Quantitative real-time PCR (RT-qPCR) analysis revealed differential regulation of these four genes in seaweeds collected during rainy and dry seasons. These EST resources provide valuable sequence information for discovering candidate genes related to important agronomic traits of *Sargassum*, and gene expression profiling analyses and functional genomics studies to elucidate their roles. The assembly and associated information provides a framework for future investigations in functional genomics in *Sargassum* species.

ABSTRAK

Sargassum polycystum C. Agardh adalah rumpai laut perang yang penting dari segi ekonomi dan ekologi, dan menjadi salah satu bahan mentah utama untuk penghasilan alginat di negara-negara Asia dengan aplikasi yang banyak dalam industri makanan, makanan ternakan, farmaseutikal dan perubatan. Walau bagaimanapun, keseluruhan genomnya tidak pernah disusun, dan hanya sumber-sumber terhad sahaja yang terdapat di GenBank untuk memahami mekanisma molekul yang berasaskan biosintesis alginat. Dalam kajian ini, 2577 `tag susunan yang diekspresikan' (ESTs) yang berkualiti tinggi dihasilkan daripada perpustakaan cDNA yang diperolehi daripada RNA terkumpul yang diasingkan daripada sampel *S. polycystum*. ESTs digabungkan dengan parameter yang ketat, menghasilkan 295 _contig' dan 1429 _singleton', dan memberi sejumlah 1724 _unigene'. Daripada semua jangkaan protein yang dikodkan oleh unigene-unigene ini, 46.3 % mempunyai padanan yang signifikasi ke atas protein yang dikenalpasti dalam pangkalan data Swiss-Prot, dan 66.2 % menunjukkan persamaan signifikasi kepada jujukan dalam pangkalan data NCBI _protein tidak berulang'. Unigene-unigene tersebut telah dicirikan fungsinya oleh anotasi Gene Ontologi (GO), di mana ~ 40.3 % dikelaskan ke dalam satu atau lebih kategori GO. Kami telah mengenal pasti dua unigene yang mengekodkan satu potensi enzim yang berkaitan dengan alginat, _mannuronan C5-epimerase' (MC5E) dan dua unigene yang mengekodkan enzim yang berkaitan dengan _sulfated-fucan', _GDP mannose 4,6 dehyratase' (GM46D), dan _GDP-4-keto-6-D-mannose-epimerase-4-reductase' (GFS), masing-masing. Analisis kuantitatif masa sebenar PCR (RT-qPCR) mendedahkan perbezaan pengaturan keempat-empat gen dalam rumpai laut yang dikumpul semasa musim hujan dan musim kering. Sumber-sumber ESTs ini memberikan maklumat urutan yang berharga untuk penemuan calon gen yang berkaitan dengan sifat-sifat agronomi penting dalam *Sargassum*, serta analisis profil ekspresi gen dan kajian genomik berfungsi untuk menjelaskan peranan mereka. Penggabungan tersebut dan maklumat yang berkaitan menyediakan satu rangka kerja untuk kajian dalam genomik berfungsi bagi spesies *Sargassum* pada masa hadapan.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank both of my supervisors Prof. Dr. Phang Siew Moi and Prof. Dr. Ho Chai Ling for their continuous support throughout my PhD with patience and knowledge. They have provided me excellent guidance, advice, and knowledge for the past several years and during the write-up of research articles and thesis.

Secondly, I would like to thank all my labmates and friends including, Dr. Yeong Hui Yin, Dr. Emienour, Dr. Yow Yoon Yen, Dr. Ng Fong Lee, Dr. Poong Sze Wan, Dr. Ng Poh Kheng, Dr. Song Sze Looi, Dr. Wong Ching Lee, Dr. Ng Woan Shien, Dr. Tan Ji, Dr. Chan Cheong Xin, Tan Cheng Yao, Fiona Keng, Pui Ling, Mei Cing, Jeannette Lai, Mindy, Kok Keong, James Lim, Muhammad, Nizam, Yung Chie, Seddon Teoh, Syin Ying, Ee Leen, Rouh San, Keat Ai, Yu Lok, Eng An, Erica Khew and Conie Toh. They have contributed immensely towards my personal and professional time in University of Malaya and University of Putra Malaysia, respectively, and created a friendly and productive reseach environment. On top of that, I would like to thank Tan Yung Chie and Teoh Seddon, who always willing to help and have provided indispensable advise and best suggestions during the write-up of this thesis.

I would like to express my gratitude to University of Malaya Fellowship Scheme, E-Science Fund from the Ministry of Science, Technology and Innovation (MOSTI) Grant (05-01-03-SF0188); University of Malaya Postgraduate Research Fund (PS248/2008C) for making my research studies possible.

Finally yet importantly, I would like to thank my parents, my husband and my son and my family members for their unconditional support and encouragement through the years, which provided me the motivation to keep going to accomplish my study. I owe them a great debt of gratitude and thus, I would like to dedicate this thesis to them as a symbol of my appreciation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

β	:	Beta
λ	:	Lambda
µg	:	Microgram
µL	:	Microliter
°C	:	degree
Δ	:	Delta
®	:	registered trademark
1 x	:	1 time
10 x	:	10 time
24:1	:	ratio 24 to 1
25:24:1	:	ratio 25 to 24 to 1
A_{260}/A_{230}	:	Absorbance at wavelength of 260 nm over 230 nm
A_{260}/A_{280}	:	Absorbance at wavelength of 260 nm over 280 nm
BLAST	:	Basic Local Alignment Search Tool
bp	:	base pairs
CAP3	:	Contig Assembly Program 3
cDNA	:	complementary DNA
cfu	:	colonies forming unit
Cl	:	Chloride
CsCl	:	cesium chloride
CTAB	:	hexacetryltrimethyl ammonium bromide
Ct	:	threshold cycle
dbEST	:	database of EST
dATP	:	2'-deoxy-adenosine-5'-triphosphate

dCTP	:	2'-deoxy-cytidine-5'-triphosphate
DEPC	:	diethyl pyrocarbonate
dGTP	:	2'-deoxy-guanosine-5'-triphosphate
DMSO	:	Dimethylsulfoxide
DNA	:	deoxyribonucleic acid
DNase	:	Deoxyribonuclease
dNTP	:	2'-deoxy-deoxyribonucleoside-5'-triphosphate
DTT	:	Dithiothreitol
dTTP	:	2'-deoxy-thymidine-5'-triphosphate
E	:	Efficiency
EDTA	:	ethylenediaminetetraacetic acid
EST	:	expressed sequence tag
EtBr	:	ethidium bromide
g	:	Gram
g	:	relative centrifugal force (*rcf*)
H_2	:	Hydrogen
HCl	:	hydrochloric acid
kb	:	kilo base pairs
L	:	Liter
LB	:	Luria-bertani
LiCl	:	lithium chloride
M	:	Molar
$MgSO_4$	:	magnesium sulfate
min	:	Minute
mL	:	Milliliter
mM	:	Milimolar

mRNA	:	messenger RNA
NA		not available
NaCl	:	sodium chloride
NaOAc	:	sodium acetate
NaOH	:	sodium hydroxide
NCBI	:	National Center for Biotechnology Information
ng	:	Nanogram
nr		non-redundant
OD_{600}	:	optical density at wavelength of 600 nm
PCR	:	polymerase chain reaction
pfu	:	plaque forming units
RNA	:	ribonucleic acid
RNase A	:	ribonuclease A
rpm	:	revolution per minute
ROS	:	reactive oxygen species
SDS	:	sodium dodecyl sulphate
sp	:	Species
TAE	:	tris-acetate EDTA
Tm	:	melting temperature
™	:	Trade mark
Tris	:	2-amino-2-hydroxymethyl-1,3-propandiol
U	:	Unit
v/v	:	volume per volume
w/v	:	weight per volume

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Background

Seaweeds are macroscopic, multicellular, marine ecosystems as they serve as shelter, nursery grounds and food for many marine organisms (Battacharyya *et al*., 2015). The use of seaweed has a long history in human civilization. The 15,000-year old human artifacts were recovered from the early archaeological site in Monte Verde, Southern Chile. It was the earliest proof that early nomadic group consume the seaweeds as food and medicine (Dillehay *et al*., 2008). Seaweed has been traditionally used as food, animal feed, fertilizer, source of phycocolloids and medicine for centuries, since it produces a wide range of metabolites such as lipids, minerals, carbohydrates (sugars), vitamins, and bioactive polyphenols, proteins (Rajauria *et al*., 2015; White & Wilson, 2015). The discovery of phycocolloids had led to the application of seaweeds in a wide variety of industries. The seaweed industry is an ever-growing business. The total production of seaweeds worldwide from mariculture and harvesting of wild stocks increase from 21 million tonnes (wet weight) on 2010 to 29.5 million tonnes (wet weight) on 2013 (FAO, 2016). The total value of worldwide seaweed industrial products for direct or indirect human uses has estimated of US$ 6 billion per year (FAO, 2014), and alginate, which is uniquely produced by the brown seaweeds (Phaeophyceae) contributed one-fifth of this value. China is the top producer of the browns seaweeds, supplying over 50 % of the world production from 2003 to 2013, and Indonesia ranked second from 2007 to 2013 (FAO, 2016).

Alginates are carbohydrates that are widely used in food industry, pharmaceutical industry, feed, paper industry, environmental and other various industries (Samaraweera *et al*., 2012; Sartal *et al*., 2012; Lee & Mooney, 2012). Presently, most of the alginate produced commercially is still obtained from brown seaweed. *Sargassum* spp. are one

of the temperate to tropical brown seaweed that are commonly used for alginate production in Southeast Asian countries (FAO, 2016). In Malaysia, *Sargassum* is abundant and 39 species have been identified (Phang *et al*., 2008), and one of the most abundant brown seaweeds is *Sargassum polycystum*. In Teluk Kemang brown seaweeds are densely populated along the reef flats, and the most conspicuous of them being *Sargassum* (Yeong & Wong, 2013). *Sargassum polycystum* is one of the dominating brown seaweed at Teluk Kemang (Yeong & Wong, 2013). Fasihuddin and Siti (1994) reported that high quality alginate obtained from *Sargassum* was comparable with commercially used seaweeds (*Macrocystis pyrifera*, *Ascophyllum nodosum* and *Laminaria hyberborea*). However, many aspects of *Sargassum polycystum* remain poorly explored, thus it has a high potential for new discoveries.

Functional genomics is a field of molecular biology that try to utilize the enormous riches of data produced by genomic projects which are genome sequencing projects, to interpret the functions of genes and proteins and their interactions (Wikipedia, 2016a). In this study, the Expressed Sequence Tags (ESTs) approach was used for the functional genomic studies on *S. polycystum*, which has very little genomic information available.

The EST approach is an efficient way to obtain a high number of genes that are expressed in an organism or tissues. EST are unedited, automatically processed, single-read sequences produced from cDNAs, and has been proven useful in many applications such as discovery of novel genes or transcripts, recovery of genomic clones and/ or full length cDNA, development of genetic maps, recognition of exons, delineation of protein families, identification of organism- or tissue- specific genes, and research on unidentified/ new functions (Rudd, 2003). The first seaweed EST database was constructed with 200 sequence tags from *Gracilaria gracilis* (red seaweed) (Lluisma & Ragan, 1997). This was followed by the generation of 10,154 ESTs from *Porphyra*

yezoensis (red seaweed) (Nikaido *et al*., 2000). As for brown seaweeds, the first ESTs was generated from *Laminaria digitata* with the sequencing of 905 ESTs (Crépineau *et al*. 2000). In 2005, Roeder and co-workers generated 1,985 ESTs from protoplasts of this alga (Roeder *et al*., 2005), followed by additional EST information acquired from several brown algae: 1,876 from *Sargassum binderi* (Wong *et al*., 2007), 90,637 from *Ectocarpus siliculosus* (Dittami *et al*., 2009), 12,000 from *Fucus serratus* and *Fucus vesiculosus* (Fucales, Ochrophyta) (Pearson *et al*., 2010), 574 from *Laminaria japonica* (Xuan *et al*., 2012) and 400,503 from *Saccharina latissima* (Heinrich *et al*., 2012). In 2014, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences has used RNA-sequencing technolgy for *de novo* transcriptome analysis of *Sargassum thunbergii* (Liu *et al*., 2014). In 2010, the genome of *Ectocarpus siliculosus*, a non-commercial brown seaweed, was sequenced and completed by the Roscoff Marine Biological Research Station and Genoscope, France (Cock *et al*., 2010). In 2015, the genome of *Saccharina japonica*, one of the most economically important brown seaweeds, has been sequenced and completed by Beijing Institute of Life Science, Chinese Academy of Sciences, China (Ye *et al*., 2015). However, the manual curation for both the genome has yet to be completed.

A good cDNA library needs to be constructed in order to generate ESTs from *Sargassum polycystum*. Construction of the cDNA library begins with the preparation of high quality total RNA. However, extraction of RNA from seaweeds has always been problematic. Thus, it is important to optimize the isolation of RNA from *S. polycystum*. Seaweeds contain high concentrations of polysaccharides. During the nucleic acid extraction, polysaccharides and secondary metabolites are released after the disruption of the cell (Chan *et al*., 2007). Well-established RNA extraction methods for higher plants or and other seaweed species may not work well for *S. polycystum*. This is because different types of seaweed may have different amount quantity and properties

of polysaccharides compared to that of higher plants and other seaweeds. Therefore, it is a necessary to optimize the RNA extraction from *S. polycystum* to obtain good quality RNA with minimum contamination.

The generated ESTs in this study will be compared with the databases at National Center for Biotechnology Information (NCBI), which will result in the identification of the useful genes, especially those involved in the pathway of alginate biosynthesis, as well as novel genes. These useful and/or novel genes are significant contributions to functional genomic studies, and will contribute towards better understanding of the biochemistry and molecular biology of *S. polycystum* and other brown seaweeds in the future.

1.2 Research Statement

Expressed sequence tags (ESTs) sequencing is a rapid and cost-efficient way to acquire transcriptome data for an organism with a large, complex and unknown genome. This research addressed the following research questions to provide useful scientific information for future functional genomics studies of *S. polycystum* and other brown seaweeds.

1.3 Research Questions

(a) What is the best total RNA extraction method for *Sargassum polycystum* in order to construct a good cDNA library?

(b) What are the gene functions of the ESTs of *Sargassum polycystum*?

(c) Are the interesting genes related to the biosynthesis of alginate and/or sulfated fucan how are they being regulated at different seasons (rainy and dry seasons, respectively)?

1.4 Objectives

The objectives of this study are:

(a) To develop and optimize a total RNA extraction method for *Sargassum polycystum*;

(b) To generate and sequence 2,500 ESTs from *Sargassum polycystum*;

(c) To identify the functions of the ESTs of *Sargassum polycystum* by comparing the sequences with the existing information in the database; and

(d) To analyse and profile the transcript abundance of ESTs related to the biosynthesis pathways which are unique to brown algae.

1.5 Research Approach

The research approach of the present study are;

Stage 1: RNA Extraction Methods of *S. polycystum*.
Method 1: Modified Hong et al. (1997)
Method 2: Modified Kim et al. (1997)
Method 3: Modified Wang et al. (2008)
Method 4: Wang et al. (2007)
Method 5: New optimized method developed in this study

↓

Stage 2: Quantification of total RNA.
Quantification of total RNA was conducted by using spectrophotometer and formaldehyde-agarose gel electerophoresis.

↓

Stage 3: mRNA isolation and quantification of mRNA.

↓

Stage 4: cDNA library construction and ESTs generation.

↓

Stage 5: Functional annotation of assembled ESTs (unigenes) from *Sargassum polycystum.*

↓

Stage 6: Sequence analysis of interesting genes which unique to biosynthesis pathway of brown algae.

↓

Stage 7: cDNA synthesis for quantification reverse transcriptase-polymerase chain reaction (RT-PCR).

↓

Stage 8: Selection of endogenous genes.
Three most stable endogenous genes (alpha tubulin, ribosomal protein L3a and translocation elongation factor 1a) from *S. polycystum* were selected for reverse transcription quantitative real-time PCR analysis.

↓

Stage 9: Reverse Transcriptase Quantitative Real-Time PCR (RT-qPCR) analysis.

↓

Stage 10: Discussion and Answer to Research Questions.

Figure 1.1: Outline of the research approach.

CHAPTER 2: LITERATURE REVIEW

2.1 Seaweeds

Seaweeds are marine macroalgae and most of them are classified as benthic, where they grow from the seabed (Mouritsen, 2013). Seaweeds usually have root-like structures called ‗holdfasts' which anchored the plant to rocks or other solid structures at the sea bottom, which generally perform the sole function of attachment (New World Encyclopedia, 2015). Seaweeds are of ecological importance where in most of the oceans, they dominate the rocky intertidal, and cover rock surfaces in the shallow subtidal in temperate and polar regions (Wiencke & Amsler, 2012).

Seaweeds are classified into three major groups, brown algae (Phaeophyceae, Ochrophyta), red algae (Rhodophyta) and green algae (Ulvophyceae, Chlorophyta), based on various properties such as pigmentation, storage contents, the organization of photosynthetic membranes, and other morphological features (Kiling *et al*., 2013). In addition that all the groups contain green pigment chlorophyll, most seaweeds possess pigments of other colors called accessory pigments. These accessory pigments give the seaweed their characteristics color, and may play a role in absorbing light of various wavelengths, and then passed on to chlorophyll *a*. The accessory pigment, xanthophyll is found in brown algae, giving them from golden brown to dark brown or black color, and it is efficient in absorbing sunlight under limited light environment (Colombo-Pallotta *et al*., 2006). The red algae contain major accessory pigments phycobilin and phycoerythin, which also aid their photosynthesis process in dim light, whereas the green algae contain carotene and chlorophyll *b* (Fong & Paul, 2011; Christaki *et al*., 2013). Storage products for Ochrophyta, Rhodophyta and Chlorophyta are different from each others. Most algae store their food starch or other products such as laminaran ((1,3)-β-D-glucan with β-(1,6) branching) and mannitol in Ochrophyta, floridean starch

(amylopectin-like) in Rhodophyta, and starch (a mixture of soluble amylase and insoluble amylopectin) in Chlorophyta (Mišurcová, 2012).

Seaweeds make up the Sargasso Sea, where two species, *Sargassum natans* (Linnaeus) Gaillon and *Sargassum fluitans* (Boergesen) Boergesen, dominate the floating algal communities of the sea (de Széchy *et al*., 2012). Some seaweeds are found to depths of 250 m in clear waters such Mediterranean, Caribbean and Brazil. The giant kelp, *Macrocystis pyrifera*, one of the largest seaweed, is commonly found in Australia, South Africa, in the southern oceans near South America and along the coast of the eastern Pacific Ocean, from Baja California north to southeast Alaska (Wikipedia, 2016b).

Seaweeds from coral reefs to kelp forests are of ecologically importance where they play a central role in coastal habitats as ecosystem engineers, competitors and primary producers (Harley *et al*., 2012). Seaweeds also are economically valuable resources where they are used as food, fodder, manure, biological treatment of wastes, as industrial raw materials and medicine, and thus useful to mankind in many ways (Ghosh *et al*., 2012).

2.1.1 Brown Algae (Phaeophyceae, Ochrophyta)

The brown algae (Phaeophyceae) are a large group of algae consisting of largely marine multicellular algae (Silberfeld *et al*. 2010). They belong to the stramenopiles or heterokonts lineage, together with organisms such as diatoms and oomycetes. Hence, they are phylogenetically distant from vascular plants, fungi and animals, and also from the green and red algae (Coelho *et al*., 2012; Katz, 2012). Many brown seaweeds inhabit the intertidal zone, in cold and temperate waters as well as in deep-water tropical

regions (Santelices, 2007). As benthic and sessile organisms, brown seaweeds have high degree of tolerance to diverse abiotic stressors such as temperature, irradiance and osmotic stress. Brown seaweeds vary from most of the land plants in many aspects of their biology, such as they can synthesize both C18 and C20 oxylipins (Ritter *et al*., 2008), have ability to accumulate iodine in their cell wall (Leblanc *et al*., 2006; Küpper *et al*., 2008; La Barre *et al*., 2010), their cell walls have unique composition, and the related cell wall biosynthesis pathway (Nyvall *et al*., 2003; Tonon *et al*., 2008; Michel *et al*., 2010b), and their use of laminarin as a storage polysaccharide (Bartsch *et al*., 2008; Michel *et al*., 2010a; Bonin *et al*., 2015).

Brown algae and other Stramenopiles exhibit some remarkable metabolic characteristic, notably in their unique carbon (C) storage metabolism (Gravot *et al*., 2010). Indeed, brown algae do not use the photoassimilate D-fructose-6-phosphate (F6D) to produce fructans, starch and saccharose as in majority of the higher plants, but alternatively produce the mannitol and β-1,3-glucan laminarin (Bonin *et al*., 2015). When most living organisms store C as linear or branched α-1,4-glucans (glycogen or starch), brown algae store C as mannitol and β-1,3-glucan (laminarin) (Michel *et al*., 2010a). The mannitol cycle involved four enzymes which are mannitol-1-phosphate dehydrogenase (M1PDH), mannitol-1-phosphatase (M1Pase), mannitol-2-dehydrogenase (M2DH), and hexokinase (HK), were identified in *E. siliculosus* (Rousvoal *et al*., 2011). Large amount of mannitol is present in brown algae and its synthesis involved two steps where a mannitol-1-phosphate 5-dehydrogenase (M1PDH) catalyzes a reversible reaction between fructose-6-phosphate (F6P) and mannitol-1-Phosphate (M1P), and then mannitol-1-phosphatase (M1Pase) hydrolyzes M1P into mannitol (Bonin *et al*., 2015). On the other hand, the putative genes encoding M1PDH, M2DH and HK were identified in *S. thunbergii* transcriptome (Liu *et al*., 2014). Liu *et al*. (2014) concluded that *S. thunbergii* probably has homologous mannitol metabolism

process as in *E. siliculosus*. Gravot *et al*. (2010) reported that mannitol content, up to 6.4 % of the dry weight according to metabolite profiling performed in *E. siliculosus*.

2.1.2 Polysaccharides of Brown Algae

Cell wall of brown algae is acidic polysaccharides which exhibit some unique characteristics, share features with both animals and plants. These polysaccharides are polymers of monosaccharides (simple sugars) which linked by glycosidic bonds (Vavilala & D'Souza, 2015). The acidic polysaccharides are mostly consisted of sulfated fucans (fucoidans) and alginic acids (alginate) or laminarins (laminarans), which has high nutritional value (Brownlee *et al*., 2009; Apostolidis & Lee, 2012), and act as promising ingredients for nutraceutical, functional, pharmaceutical and cosmeceutical industries (Dettmar *et al*., 2011; Lee & Mooney, 2012). In mature intertidal brown algae, average weight ratio of alginates, fucoidans and cellulose are 3: 1: 1, respectively (Michel *et al*., 2010b). The principal polysaccharides for Rhodophyta are carrageenans, xylans, floridean starch (amylopectin like glucan) whereas those in Chlorophyta are ulvans, sulfated galactans, xylans, and sulfuric acid polysaccharides (Ngo & Kim, 2013; Vavilala & D'Souza, 2015).

2.1.2.1 Mannuronan C5-epimerase (MC5E) Enzyme Involved in Alginate Biosynthesis Pathway

Alginate which harvests from brown algae, is industrially and commercially important polysaccharide. It is the major matrix component of brown algal cell wall, which weight up to 45 % in dry weight. It consists of α-1,4-L-guluronic acid (G) and β-1,4-D-mannuronic acid (M) residues, where G residues and M residues are distributed

as MM blocks, MG blocks and/or GG blocks (Draget *et al*., 2005). These features vary from the the growing area (exposed or sheltered), the season, and the age of the tissues (Nyvall *et al*., 2003).

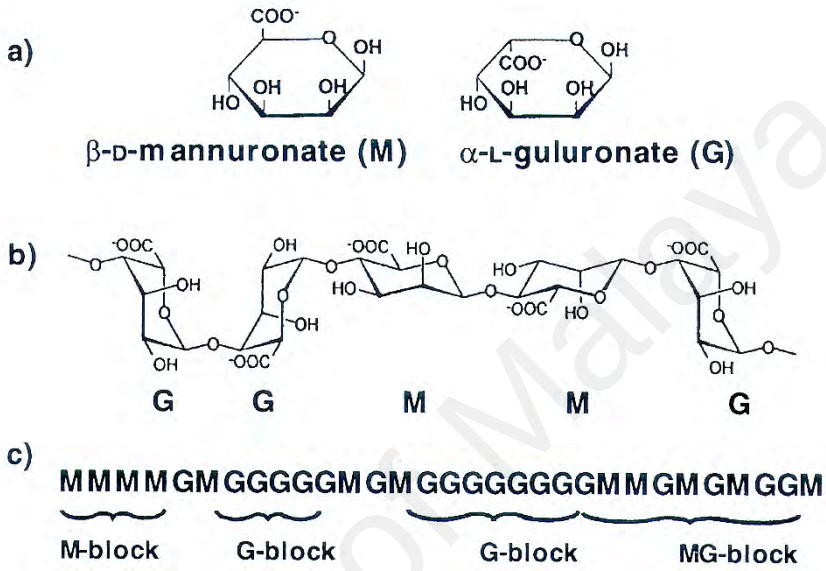


Figure 2.1: Structural characteristic of alginates: (a) alginate monomers, (b) chain conformation, (c) block distribution.

(adapted from Draget *et al*., 2005)

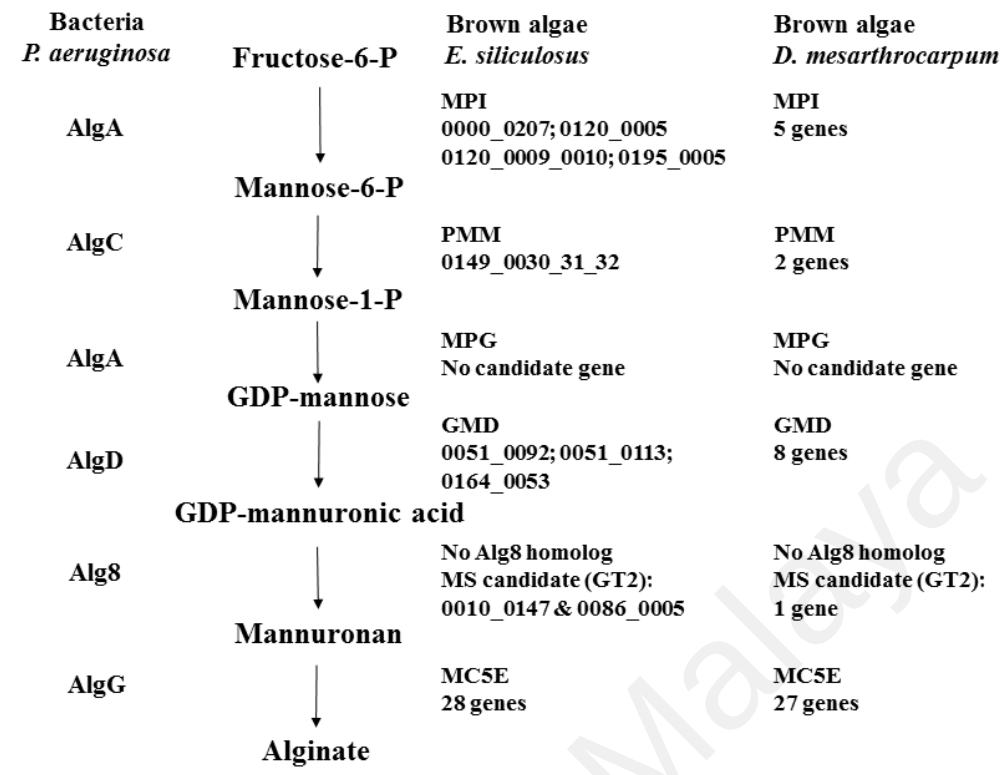
Besides alginate biological resources are mainly extracted from brown algae, it also could be obtained from opportunistic pathogens (such as, two genera of bacteria *Pseudomonas* and *Azotobacter*). *Pseudomanas aeruginosa* secretes alginate when it infects lung tissues in cystic fibrosis patients, and the viscous matrix it forms exacerbates the pulmonary difficulties suffered by the patients, and thus causes the morbidity and mortality in cystic fibrosis patients (Lyczak *et al*., 2002; Ramsey & Wozniak, 2005). Meanwhile, *Azobacter velandii* produces alginate as an extracellular polysaccharides in a process call cyst formation to protect the organism from the damaging effects of harmful conditions in the environment, for example, *A. velandii*

tends to built alginate capsules with different composition according to the external dissolved oxygen concentration, implying that its production may contribute to a protection of nitrogenase against oxygen damage (Sabra & Zeng, 2009). The presence of O-acetyl groups at C2 and/or C3 in the bacterial alginates is the main dissimilarity between bacterial and algal alginates at the molecular level (Skjåk-Braek *et al*., 1986).

Mannuronan C5-epimerase (MC5E) is the key enzyme in the final step in alginate biosynthesis, which decides the M/G ratio of alginate. The reaction is catalyzed by the enzyme MC5E which is the epimerization of β-D-mannuronic (M) residues into α-L-guluronic (G) residues within the polymer chain (Wang *et al*., 2014b) (Figure 2.2). The knowledge about the genes involved in the alginate biosynthesis mostly comes from bacterial *Azotobacter* and *Pseudomonas* (Tipton, 2010). However, in brown algae, the studies are mostly focuses on the genome sequencing analysis and transcriptome (Cock *et al*., 2010; Dittami *et al*., 2009; Liu *et al*., 2014). The first study on the isolation and characterization of cDNA encoding MC5E from brown alga, *L. digitata* was reported by Nyvall *et al*. (2003). Six different cDNAs encoding MC5E were identified among the ESTs of *L. digitata*. This was followed by 22 transcripts encoding for different MC5E identified from *L. digitata* cDNA library constructed from protoplasts (Roeder *et al*., 2005). A total of 26 MC5E genes were annotated in the *E. siliculosus* genome (Cock *et al*., 2010). Wang *et al*. (2014b) predicted a total of 94 open reading frame sequences for MC5E from brown algae from the 1000 Plant Transcriptome Sequencing Project (OneKP). They studied the conserved motifs of MC5E protein, the structure modeling of the MC5E, the algal MC5E gene family characteristics, and as well as the variation incidents that most probably would happened during evolution in algae, by using method of transcriptome sequence analysis, multiple sequence alignment and phylogenetic tree construction, and also gene structure and gene localization analysis. Liu *et al.* (2014) and Yamagishi *et al*. (2014) discovered quite a number of genes

encoding four putative enzymes which are involved in biosynthetic pathway of alginate (mannose-6-phosphate (MPI), phosphomannomutase (PMM), GDP-mannose 6-dehydrogenase (GMD) and MC5E). Nyvall *et al*. (2003) and Wang *et al*. (2014b) proposed that brown algae cDNAs encoding MC5E or the whole alginate biosynthesis pathway would have arisen from ancient bacteria (*P. aeruginosa* and *A. vinelandii*) via horizontal gene transfer and subsequently it diverged into gene family during evolution. Wang *et al*. (2014b) reported that bacterial and brown algae MC5E gene might belong to the same structural family and arisen from a common ancestor, which they display a similar structure and share the same catalytic domain.

The expression of genes encoding MC5E, reported by Nyvall *et al*. (2003), was up-regulated in spring and winter. Roeder *et al*. (2005) found that genes encoding mannuronan C5-epimerase are highly expressed in *L. digitata* protoplasts. In the *Ectocarpus* genome, genes encoding mannuronan C5-epimerase were highly represented, and under different stress conditions, some of them were either repressed or induced. However, none of them were among the most significantly regulated loci (Dittami *et al*., 2009). Tonon *et al*. (2008) supported the hypothesis that under pathogen attack, specific MC5E genes are obtained, based on the result of transcript accumulation during sporophyte elicitation and protoplast regeneration, which in response to environmental stimuli, *L. digitata* enable to modify its cell wall. Wang *et al*. (2014a) used the Reads per kilobase of gene model per million mapped reads (RPKM) to quantitate the global gene expression level of six *Sargassum* species (*S. horneri, S. vachellianum, S. hemiphyllum, S. henslowianum*, *S. fusiforme,* and *S. integerrimum*) China. The results showed that all the six species have the alginate transcripts, but in different expressed numbers.



The alginate-related enzymes from *P. aeruginosa* are indicated on the left-hand side. Gene products which are likely to belong in the alginate biosynthesis pathway in *D. mesarthrocarpum* and *E. siliculosus* are indicated on the right-hand side by the code XXXX_YYYY, where XXXX indicates the supercontig. The prefix Esi has been omitted for clarity. MPI, mannose-6-phosphate isomerase; PMM, phosphomannomutase; MPG, mannose-1-phosphate guanyltransferase; GMD, GDP-mannose 6-dehydrogenase; MS, mannuronan synthase; MC5E, mannuronate C5-epimerase.

Figure 2.2: The brown algae (*Ectocarpus siliculosus* and *Discosporangium mesarthrocarpum*) and bacteria (*Pseudomonas aeruginosa*) alginate biosynthesis pathway.

(adapted from Michel *et al*., 2010b; Yamagishi *et al*., 2014)

In the co-polymer, the amounts of three types of blocks (G, MG and M) are mainly determined by the physical properties of alginates (Brownlee *et al*., 2009). In general, the elastic and soft alginate gel was produced by alginate with a high M/G ratio and a low amount of guluronic blocks, whereas a rigid and strong alginate gel was formed by a low M/G ratio and a large amount of guluronic blocks (Bourgougnon & Stiger-Pouvreau, 2012). The stiffness of the chain blocks increased in the order MG < MM < GG was determined by the viscosity data of alginate solutions. *Laminaria hyperborea*

grows in an exposed coastal area, thus a high content of guluronic acid was found in the holdfast and the stipe, which are stiffer than its blades. Whereas the blades of *L. hyperborea* containing lower guluronic acid content, exhibited a more flexible texture. Alginates with low gel strength and low content of G-blocks were generally prepared from *Macrocystis pyrifera*, *Ascophyllum nodosum* and *Laminaria japonica* whereas alginate with high contents and high gel strength of G-blocks were extracted from stipes of old *L. hyperborea* (Draget *et al*., 2005).

2.1.2.2 Enzymes GDP-Mannose 4,6-Dehydratase (GM46D), GDP-4-Keto-6-Deoxy-D-mannose Epimerase/Reductase (GFS), Fucosyltransferases, and Sulfotransferases (STs) Involved in Sulfated Fucan Biosynthesis Pathway

Algal sulfated galactans (SGs) and sulfated fucans (SFs) are the most studied sulfated polysaccharides from the marine orgnisms (Pomin, 2012). Brown algal SFs are in generally more intricate and heterogenous than red or green algal SGs. Controversial published data by different authors occurred when the structures of these brown algae are hard to determine (Pomin, 2012). The complexity of the systems is increased by random distribution of sulfation, the presence of branching residues, and also the number of types, positions of glycosidic bonds and the presence of other heterogeneities such as acetylation, methylation, and pyruvylation (Pomin, 2012; Bilan *et al*., 2013). Basically, SFs, polysaccharides that contain neutral sugars, substantial percentages of L-fucose and sulfated ester groups, are occurring in brown seaweed cell walls (Phaeophyceae, mainly Fucales, Laminariales and Ectocarpales) (Dore *et al*., 2013), and also in the jelly coat of sea urchin egg and body wall of sea cucumbers (Michel *et al*., 2010b). In the first report of isolation of an SFs in 1913 from marine brown algae, fucoidin was the initial name for the fucose containing glycan (Pomin, 2012). In 1967,

McDowell renamed it as fucoidan (Pomin, 2012). Now it is names as –SFs" in agreement with IUPAC recommendations, which defines SFs as polysaccharides consisting of at least 90 % sulfated L-fucose by monosaccharide composition. This term was applied to the sulfated fucans of marine invertebrates, whereas the term fucoidan or algal sulfated fucans have been used for fucans extracted from algae. For clarity, and because the fucans extracted from those two sources are different, these two nomenclature will be adopted in my study.

Fucoidans from brown seaweed consist of backbone of (1→3)- and (1→4)-linked α-L-fucopyranose residues, which may be organized in stretches of (1→3)-α-fucan or of alternating α(1→3)- and α(1→4)-linked L-fucopyranose residues (Ale *et al*., 2011). These backbones are usually substituted with sulfate or acetate, and/or have side branches containing fucopyranoses or other glycosyl units, for example glucuronic acid (Ale & Meyer, 2013). Several of the fucoidan structures reported in the literature have also contained small amounts of various other monosaccharide residues, e.g. glucose, galactose, xylose, rhamnose, and/or mannose, giving rise to extremely complex heterogeneous mixtures of molecules (Ale *et al*., 2011; Huang *et al*., 2016). Besides, other sulfated fucose-rich heteropolysaccharides composed of many other sugars such as ascophyllan, xylofucoglycuronan, sargassan, glycuronofucoglycan and sulfated galactofucan also have been characterized and extracted from marine brown algae (Pomin, 2012; Bilan *et al*., 2013; Ale & Meyer, 2013; Mohsin *et al*., 2014). Ehrig and Alban (2015) reported that sulfated polysaccharides of *Saccharina latissima* had the highest yield with sulfated galactofucan as the main fraction (67 %), where the *S. latissima* growing at an intertidal zone with high salinity and harvested at end of the growing period. The metabolism of the sulfated fucans in brown algae has yet been characterized (Michel *et al*., 2010b).

Algal SFs is located in the cell wall of brown seaweeds from where they can be extracted by using acidic solutions (Lorbeer *et al*., 2015), hot water (Hahn *et al*., 2012), or proteolytic digestions (Marques *et al*., 2012). Some studies have shown a correlation between fucoidan content and the depth at brown algae grow; the closer the algae to the surface of seawater, the greater the fucoidan content are (Berteau & Mulloy, 2003).

Fucoidans have been isolated and studied from different species of brown algae, due to their interesting and varied biological activities, including gastric protective effects, antithrombotic, anticoagulant, antioxidant, anti-inflammatory, antiviral, antibacterial, activity against uropathy, renalpathy and hepatopathy, therapeutic potential in surgery and as prebiotics (Pielesz & Biniaś, 2010; Holdt & Kraan, 2011; Pomin, 2012; Wijesinghe & Jeon, 2012; Dore *et al*., 2013; Cao *et al*., 2014; Rupérez *et al*., 2014).

In both eukaryotes and prokaryotes, L-fucose is a vital simple sugar (monosaccharide). It is a sugar component of cell wall polysaccharides in bacterial lipopolysaccharides, fungi, and plants (Ren *et al*., 2010). Glycolipids and glycoproteins play crucial part especially in biological roles such as membrane stabilization, immune response and protection, and recognition and signaling are included in the glycoconjugates (Ren *et al*., 2010).

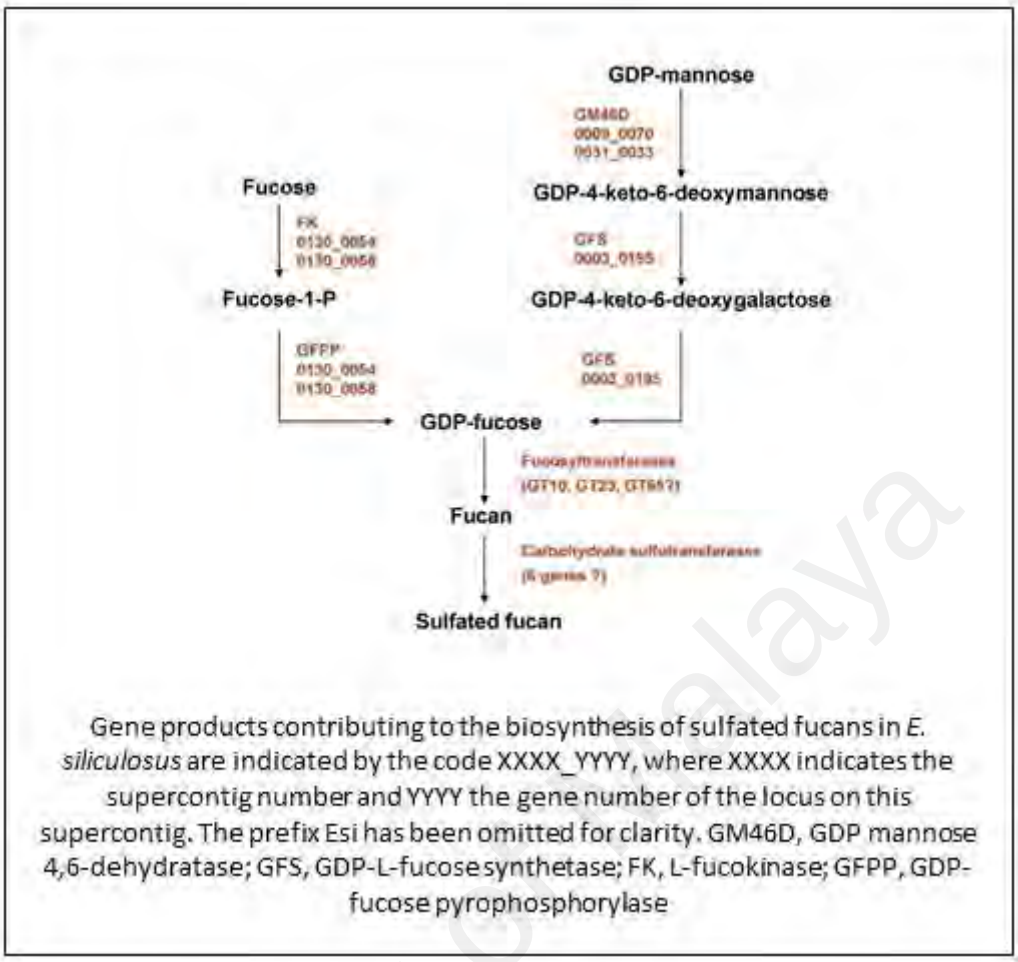


Figure 2.3: The sulfated fucans biosynthetic pathway.

(adapted from Michel *et al*., 2010b)

Via the *de novo* pathway in mammals, plants and most of the bacteria, most of the time GDP-L-fucose is synthesized from GDP-D-mannose. GDP-D-mannose, catalyzed by the GDP-mannose 4,6-dehydratase (GM46D) to GDP-4-keto-6-deoxymannose and then to GDP-fucose by GDP-L-fucose synthetase (GFS), a single enzyme which catalyzes both epimerisation and reduction (Figure 2.3). GFS also has been identified as GDP-4-keto-6-deoxymannose 3,5-epimerase/4-reductase (Michel *et al*., 2010b; Ren *et al*., 2010). An alternative salvage pathway, L-fucose is converted to GDP-fucose by enzymes fucose kinase (FK) and subsequently by GDP fucose pyrophosphorylase (GFPP) (Figure 2.3), is mostly found in animals and plants (Ren *et al*., 2010). Michel *et*

al. (2010b) reported that *Ectocarpus* harbors one GFS and two GM46D which are more far apart from the plant enzymes (30 % for GFS and 50 % for GMD), but are vastly comparable to their fungal and animal counterparts (65 % sequence identity). These proteins reveal that GDP-fucose in brown algae can also be generated via the *de novo* pathway (Michel *et al*., 2010b). One or several fucosyltransferases polymerized sulfated fucans as neutral polysaccharides, and subsequently by certain sulfotransferases (STs). *Ectocarpus* harbors four fucosyltransferases, which are glycosyltransferases (GTs): GT65, GT10 and GT23 and 15 different genes encoding STs (Michel *et al*., 2010b). However, the GT10, GT23 and GT65 are enzymes which commonly associated with protein glycosylation in other eukaryotic phyla. Thus, it is difficult to predict GTs in brown algae are novel which cannot be identified via sequence comparisons or whether they are genuine fucan synthases (Michel *et al*., 2010b). On the other hand, the need for sulfation on the fucan backbone at different positions was probably reflected by the multiplicity of these enzymes.

2.1.3 Utilisation of Brown Algae

2.1.3.1 The industrial values and the applications of alginate in industries

Brown seaweeds (Phaeophyceae) are widely used as a source of alginate (Bourgougnon & Stiger-Pouvreau, 2012). Alginates occur as capsular polysaccharides in soil bacteria. They comprise up to 40 % of the dry matter as a structural component in marine brown seaweeds. All commercial alginates to date are extracted from algal sources (Skjåk-Braek & Draget, 2012). In 2009, the estimated volume of sales of alginate is 26,500 tonnes with the market value of about US$ 318 million (Bixler & Porse, 2011). The Market and Markets (2015) estimated that the market for alginates

and their derivatives will be worth US$ 409.2 million by 2019, and the Asia-Pacific region is one of the prominent producers and exporters of alginates.

The cost of alginates are variable depending on the level of purity. Ordinary purified-grade alginate costs approximately US$ 12 per kilogram, low-purity and technical grade alginate approximately costs US$ 1 per kilogram, whereas sodium-alginate of 99 % food additive typically costs around US$ 400-500 per kilogram (Bixler & Porse, 2011; Alibaba, 2016).

Usually alginates are obtained from brown seaweeds belonging to the order Laminariales, Fucales and Durvillaeales. In the temperate countries, alginates are mainly extracted from *Ascophyllum nodosum*, *Durvillaea potatorum*, *Macrocystis* spp., *Saccharina* spp., *Lessonia* spp. and *Laminaria* spp., whereas in the Southeast Asian countries, alginates are mainly extracted from *Sargassum* and *Turbinaria* species (White & Wilson, 2015). In India, the annual production of alginates is about 300 tonnes extracted from 3000 tonnes of dry *Sargassum* (*S. wightii*, *S. ilicifolium*, *S. myriocystum* and *S. polycystum*) and *Turbinaria* (*Turbinaria conoides, T. ornate* and *T. decurrens*) harvested from wild populations (Kaliaperumal *et al*., 2004; Rao & Mantri, 2006). The yield of alginate from *S. polycystum* was reported to be higher than other brown seaweeds such as *S. vulgare*, *S. tenerrimum* (Saraswathi & Rengasamy, 2003), *S. siliquosum*, *Padina boryana* and *Dictyota dichotoma* (Goh, 2008).

Alginate extracted from *Sargassum* spp. from Vietnam, India, Thailand, Phillipines and Indonesia, are used mainly in the textile industry (OMRI, 2015; White & Wilson, 2015). High quality alginate which isolated from *Sargassum* was comparable to commercially used seaweeds (Fasihuddin & Siti, 1994) and thus, *Sargassum* spp. have become as one of the commercial sources of alginate in Southeast Asia. However, Wong and Phang (2004) reported that the biomass yield for the wild populations of two

dominant *Sargassum* species in the west coast of Peninsular Malaysia was too low to sustain harvesting for an alginate industry. As with other phycocolloids, various grades of alginate are available for specific applications and the associated prices are different; e.g. sodium alginate of pharmaceutical grade, food grade, cosmetic grade, textile printing grade, and welding rod flux grade (Bixler & Porse, 2011). Alginates are mainly used in the textile (50 %) and food (30 %) industries. Table 2.1 shows some of the main applications of alginate in various industries.

Table 2.1: The applications of alginate in industries

(Bixler & Porse, 2011; Parapurath *et al*., 2012; OMRI, 2015).

Property	Product	Application
Water-holding	Paper coating	Control rheology of coatings; prevents dilatancy at high shear
	Paper sizing	Improves surface properties, ink acceptance, and smoothness.
	Adhensives	Controls penetration to improve adhensive and applications.
	Textile printing	Produces very fine line prints with good defination and excellent washout.
	Textile dyeing	Prevents migration of dyestuffs in pad dyeing operations. (Alginate also compatible with most fibre-reactive dyes.)
Gelling	Air freshener gel	Firm, stable gel produces from cold-water systems.
	Hydro-mulching	Holds mulch to incline surfaces; promotes seed germination.
	Toys	Safe, nontoxic materials are made for impressions or putty-like compounds.
	Boiler compounds	Produces soft, voluminious flocs easily separated from boiler water.
Emulsifying	Polishes	Emulsifies oils and suspends solids.
	Antifoams	Emulsifies and stabilizes the product.
	Latices	Stabilizes latex emulsions, provides viscosity.
Stabilizing	Ceramics	Imparts plasticity and suspends solids.
	Welding rods	Improve extrusion characteristics and strength.
	Cleaner	Suspends and stabilizes insoluble solids.

2.1.3.2 Applications of alginate in food

Alginates are being used as food coating, gelling agents, encapsulation and immobilisation, and as stabilising, emulsifying and thickening agents in food (Brownlee *et al*., 2009). Alginates have gel forming ability; they enhance viscocity; stabilize aqueous mixtures, and emulsions. Table 2.2 shows some of the applications in food. Toxicological studies have proved that alginates are safe for use in food, thus the Food and Drug Administration (FDA) has granted generally recognized as safe (GRAS) status to alginates.

Table 2.2: The applications of alginate in food

(Brownlee *et al*., 2009; Cofrades *et al*., 2012; Menon, 2012; OMRI, 2015).

Property	Product	Application
Water-holding	Frozen foods	Maintains texture during freeze-thaw cycle.
	Relish	Stabilizes brine, allowing uniform filling.
	Syrups	Suspends solids, controls pouring consistency.
	Frozen desserts	Provides heat-shock protection, improved flavor release, and superior meltdown.
	Dry mixes	Quickly absorbs water or milk in reconstitution.
	Bakery icings	Counteracts stickiness and cracking.
	Pastry fillings	Produces smooth, soft texture and body.
Emulsifying	Salad and dressings	Emulsifies and stabilizes the salad.
	Meat and flavor sauces	Emulsifies oil and suspends solids.
Stabilizing	Beer	Maintains beer foam under adverse conditions.
	Fountain syrups, topping	Suspends solids, produce uniform body.
	Sauces and gravies	Thickens and stabilizes for a broad range of applications.
	Milkshakes	Controls overrun and provides smooth, creamy body.
Gelling	Instant puddings	Produces firm pudding with excellent body and texture; better flavour release.
	Pie and pastry fillings	Cold water gel base for instant bakery jellies and instant lemon pie fillings. Develops soft gel body with broad temperature tolerance; improves flavor release.
	Dessert gel	Produces clear, firm, quick-setting gels with hot or cold water.
	Cooked pudding	Stabilizes pudding system, firms body and reduces weeping.
	Chiffons	Provides tender gel body that stabilizes instant (cold make-up) chiffons.

Japan, China and Republic of Korea are the largest consumers of seaweeds. In these countries, *Undaria* spp., *Scytosiphon lomentaria*, *Sargassum* spp., *Saccharina* spp., *Pelvetia siliquosa*, *Sargassum fusiforme*, *Ecklonia* spp. are brown seaweeds which mostly eaten as vegetables (White & Wilson, 2015). The breakdown of alginate market segments by volume is shown in Table 2.3. Growth of alginate markets is a mix of geographic re-distribution of established markets, such as textile growth in China, and a few new developments in food and pharmaceutical area.

Table 2.3: Alginate market segmentation over the last decade (adapted from Bixler & Porse, 2011).

Market segment	1999 volume (metric ton)	2009 volume (metric ton)
Technical grades	11,000	11,000
Food/ pharmaceutical	8,000	13,000
Animal feed	4,000	1,000
Propylene glycol alginate (PGA)	2,000	1,500
Total	25,000	26,500

2.1.3.3 Use in the medical and pharmaceutical industries

Alginates have been greatly studied with great potential use and used in numerous biomedical and clinical applications, due to its versatile, ease of gelation, comparatively low cost, low toxicity and biocompatibility (Lee & Mooney, 2012). Alginates have been used for decades in human related health applications. Alginate hydrogels are extremely versatile and adaptable biomaterial, with great potential for use in biomedical applications, including tissue engineering, drug delivery and wound healing applications. Hydrogel is a 3-D hydrophilic, cross-linked polymeric networks produced by one or more monomers simple reaction, and it is a polymeric material that mostly constituted water within its structure, and it does not dissolve in water (Ahmed, 2015).

Traditional wound dressings, such as gauze, provide only a few functions, which prevent the entry of pathogen into the wound and allow the evaporation of wound exudates by keeping the wound dry. In contrast, modern dressings, such as alginate dressings, facilitate wound healing by providing a moist wound environment (Boateng *et al*., 2008). The alginate-based wound dressings offer numerous beneficial functions to the treatment of chronic and acute wounds, where it absorbs wound fluid resulting in a hydrophilic gel over the wound which provides a good moist environment for healing and reduce bacterial wound infections. Besides, these functions also stimulate rapid healing, rapid epithelialization and granulation tissue formation (Lee & Mooney, 2012). Various alginate dressings including CovaWound™ (Covalon Technologies), 3M™Tegaderm™ (3M Healthcare), Algicell™ (Derma Sciences), AlgiSite M™ (Smith & Nephew), Sorbalgon® (Hartmann USA) and Kaltostat™ (ConvaTec) (WoundSource, 2016) are in market.

Network of blood vessels are critical for removal of metabolic waste products, trafficking of stem and progenitor cells, and transport of oxygen and nutrients to all tissues, which are critical for wound repair in the adult. In tissue engineering, alginate gels have been widely studied and exploited as a delivery vehicle of various angiogenic molecules (recombinant proteins or genes) (Lee & Mooney, 2012). Thus, entrapment of cells within alginate gel has become one of the most widely used techniques for the immobilization of living cells (Farbo *et al*., 2016). One of the first applications of alginate hydrogels in tissue engineering is the transplantation of encapsulated pancreatic islets to cure type I diabetes (Qi, 2014). In 2013, Erro and co-authors reported that 3-D alginate matrices have demonstrated great potential in the development of a bioartificial liver. Cells from human derived liver cell line (HepG2) cultured in an optimized form, was hypothesized by Erro *et al*. (2013), could be used to temporarily improve the function of deficient liver to allow enough time for its repair or for transplantation.

Alginate plays a remarkable part in controlled-release drug products from solid, hydrophilix matrix dosage forms, while thickening, gel forming, and stabilizing agents are the conventional role of alginate in pharmaceutical industry. Their extracellular alginate matrix hydrates on contact with gastric fluid after the ingestion, and the surface polymer swells to form a viscous ‗gel layer' (Dettmar *et al*., 2011). This gelatinous layer acts as a barrier to drug release as it retards water influx and drug efflux, and prevents matrix disintegration. The M/G ratio and molecular weight of the alginate can control the speed of hydration and release of the drug (Dettmar *et al*., 2011). Alginate gels have been exploited for the delivery of low molecular weight drugs (Maiti *et al*., 2009; Elnashar *et al*., 2010), hydrophobic drugs (Hazra *et al*., 2015), and also drug delivery in combination with chitosan (Delmar & Bianco-Peled, 2016).

In China, *Laminaria, Sargassum* and *Ecklonia* are commonly used as traditional medicine. *Sargassum* spp. have been used for treatment of goiter, scrofula, edema, urinary diseases and dropsy (Liu *et al*., 2012). A perusal of literature shows that the brown algae contain essential minerals, vitamins, free amino acids, mannitol, glucitols, sulphated polysaccharides and phlorotannins (Pádua *et al*., 2015). These compounds in brown algae are found to have wide spectrum of biological properties. Fucoidan (Prabu *et al*., 2013), sterol (Matloub & Awad, 2012), tannin (Kanimozhi *et al*., 2015), mannitol (Groisillier *et al*., 2014), carotenoids and chlorophyll (Seenivasan *et al*., 2012) are some of the potential antibiotic compounds in *Sargassum* spp. Raghavendran *et al*. (2006) found that the defensive action of *S. polycystum* may due to their anti-hepatotoxic and antioxidant properties against the acetaminophen-induced severe impairment in the drug metabolizing microsomal enzyme system and inhibitory action on TNF-α. This *S. polycystum* extract also provided protection against acetaminophen-induced alterations in rat liver cell structural integrity during toxic hepatitis. Rattaya *et al*. (2015) reported that methanolic extract of *S. polycystum* at 500 mg/L concentration, could effectively

inhibit the growth of *Staphylococcus aureus*, thus, can be a good source of natural antioxidant. Heo *et al*. (2008) reported fucoxanthin isolated from *Sargassum siliquastrum* possesses prominent antioxidant activity against H_2O_2-mediated cell damage which might be a potential therapeutic agent for treatment or prevention of several diseases associated with oxidative stress.

2.1.3.4 Animal feeds

For a long time, animals such as sheeps, cattles and horses that lived in coastal areas have been eating seaweed, especially in European countries where large brown seaweeds are washed ashore. In Norway, seaweed meal has been produced for animal feeds since 1960s, using *A. nodosum*. France has used *L. digitata*, Iceland has used both *Ascophyllum* and *Laminaria*, whereas United Kingdom has used *Ascophyllum*. These seaweed meals have increased the fertility and birth-rate of their animals. This is due to fucoxanthin, a pigment in brown seaweeds, which increases the iodine content of the eggs. An experiment for seven years with dairy cows (seven pairs of identical twins) showed an average increase in milk production of 6.8 % that lead to 13 % more income. Another trial involving two groups each of 900 ewes showed that those fed with seaweed meal for more than two-year period, maintained their weight much better during winter feeding and also gave greater wool production (McHugh, 2003). Marin *et al*. (2009) reported that *Sargassum* spp. can be used as feed supplement for sheeps, especially in tropical and subtropical regions, where *Sargassum* spp. are available. *Sargassum* also can be used as fodder for sea cucumbers, urchins, abalones and winkles (Yu *et al*., 2014). *Sargassum* spp. are used as fertilizers and soil conditioners (Kumari *et al.*, 2013).

2.1.3.5 Bioremediation

The increased use of metals and chemicals in the processing industries has resulted in the generation of large quantities of aqueous effluents that contain high levels of heavy metals, thus creating serious environmental and disposal problems (Opeolu *et al*., 2010). A number of methods have been used for the removal of heavy metal pollutants from wastewaters, including precipitation, evaporation, electroplating, ion-exchange and membrane processes (Abbas *et al*., 2014). These processes are expensive, and also have other shortcomings, such as incomplete removal of metals, limited tolerance to pH change, moderate or no metal selectivity, and production of toxic sludge or other waste products that lead to further disposal problems (Abbas *et al*., 2014). Therefore, biosorption is a promising alternative method to treat industrial effluents, mainly because of its low cost and high metal binding capacity.

The cell wall of brown seaweed *Sargassum* spp. constituted of mainly alginates, usually calcium and sodium alginates, thus with a high potential for the accumulation of heavy metals (Tabaraki & Nateghi, 2014; Sargin *et al*., 2016). These polysaccharides are produced due to the interaction between alginic acid and alkaline and alkaline-earth elements from seawater. Those elements constitute efficient ion-exchangers for heavy metals present in solution. On the other hand, calcium, when in solution, is usually precipitated as calcium salts, performed during primary effluent treatment. Marta *et al*. (2006) reported that *S. filipendula* was effectively used in the biosorption of radioactive thorium. Padilha *et al*. (2004) reported that the biomass of *Sargassum* sp. is useful for the recovery of ionic copper from highly concentrated solutions simulating effluents from semiconductor production. *Sargassum* spp. have been used as a bioindicator to remove pollutants from wastewater and classified as a good biosorbent (Lodeiro *et al*., 2006; Saravanan *et al*., 2011). Jothinayagi and Anbazhagan (2009) reported that *Sargassum wightii* was effectively used as bioaccumulation study from the

polluted sites. *Sargassum* spp. also stabilize shorelines against wave erosion and provide food, shelter and nursery space for important fishery species. *S. hemiphylllum* and *S. henslowianum* have the potential to be used for marine pollution bioremediation in the coastal areas of South China, especially in mariculture zones (Yu *et al*., 2014).

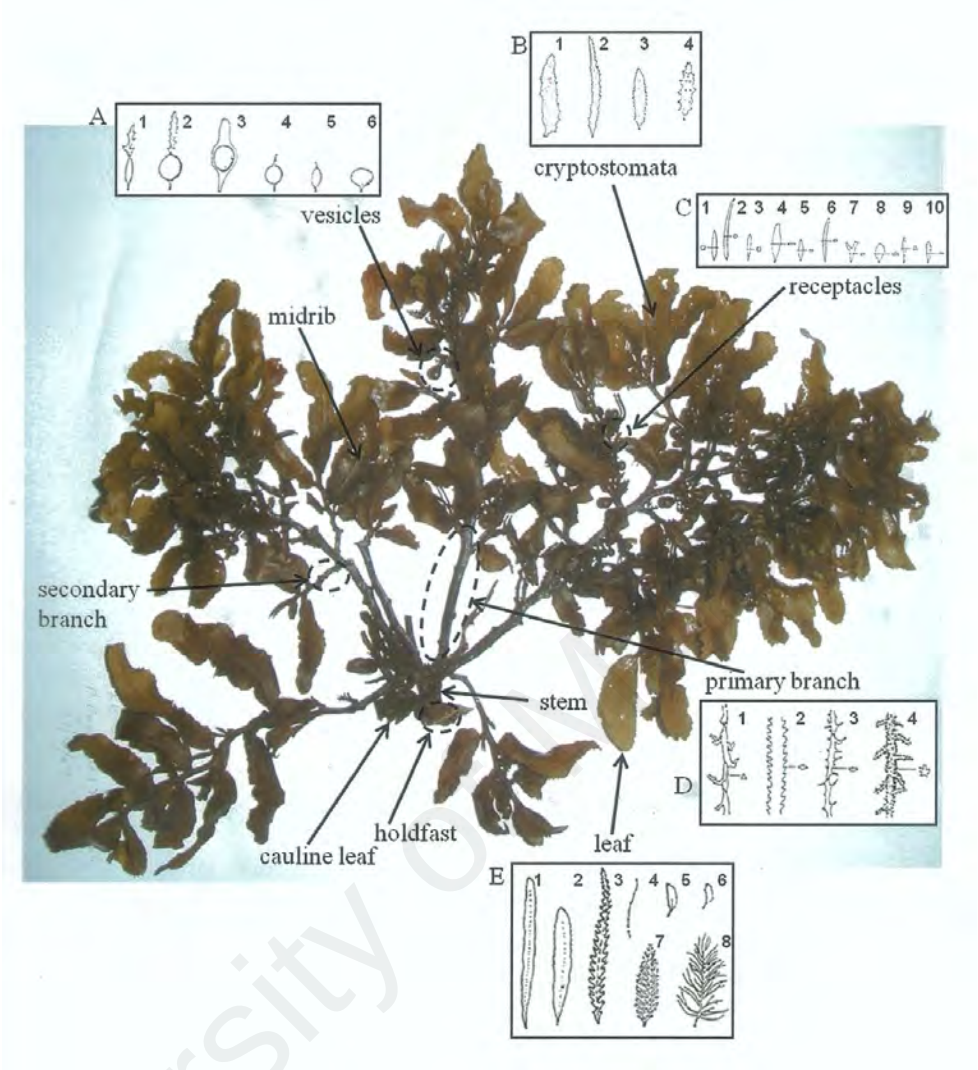
2.1.4 *Sargassum* C. Agardh

Sargassum C. Agardh 1820 (*nom. cons.*) is one of the largest genera in the Phaeophyceae, which are distributed mostly on rocky coast in temperate (Komatsu *et al*., 2014) and tropical regions (Ang, 2007) globally, and the morphology of *Sargassum* are the most complex and differentiated in the Fucales, family of Sargassaceae (Yoshida, 1989). The genus *Sargassum* was established by C. Agardh in 1820 based on morphological features such as the development of axes, shape of leaves, vesicles and receptacles. It was typified by *Sargassum bacciferum* (Turner) C. Agardh (Yoshida, 1983), and it is known as *S. natans* (Linnaeus) Boergesen currently which is common species in the Sargasso Sea in the central Atlantic Ocean (Yoshida, 1989). The classification of the taxonomy of *Sargassum*: Empire: *Eukaryota*; Kingdom: *Chromista*; Phylum: *Ochrophyta*; Class: *Phaeophyceae* Kjellman; Subclass: *Fucophycidae* Cavalier-Smith; Order: *Fucales* Bory; Family: *Sargassaceae* Kützing; Genus: *Sargassum* C. Agardh (Guiry & Guiry, 2016).

The classification system of C. Agardh was revised by J. G. Agardh and divided the genus *Sargassum* into five subgenera. After several modifications, currently, four subgenera are recognized: *Sargassum* subg. *Arthrophycus* J. Agardh; subg. *Bactrophycus* J. Agardh; subg. *Sargassum*, and sub. *Phyllotrichia* (Areschoug) J. Agardh (Womersley, 1954; Yoshida *et al*., 2004; Mattio *et al*., 2010). To date, there are 941 species names in the database, of which 353 have been flagged as currently

accepted taxonomically (Guiry & Guiry, 2016). As of January 2016, nucleotide sequence data are available at Genbank (http://www.ncbi.nlm.nih.gov/Genbank) for 2,731 samples identified as *Sargassum*. There are 39 species of *Sargassum* in Malaysia (Phang *et al*., 2008).

Sargassum plants consist of holdfast, stem, branch, leaves, leaves margin, cryptostomata, vesicles and receptacles (Phang *et al*., 2008). The descriptions of the morphology of a *Sargassum* plant is shown in Figure 2.4. The *Sargassum* has two methods of reproduction, where some species are capable of vegetative regeneration by fragmentation, while some species use the usual reproduction method which is sexual and is oogamous. Figure 2.4 shows the morphology of a *Sargassum* plant**.**



(A = variation in vesicles morphology; B = distribution of cryptostomata; C = variation in receptacle morphology; D = various type of main branch; E = various types of leaves) (adapted from Ng, 2013)

Figure 2.4: The morphology of a *Sargassum* plant.

(adapted from Ng, 2013)

Caption for Figure 2.4 with description mentioned in Section 2.1.4.

A ([1] elliptical apiculatus; [2] mucronatus with coronal leaf; [3] phylocyst; [4] ovatus; [5] ellipticus; [6] obovatus); B ([1] along leaves margins; [2] along midrib; [3] scattered; [4] prominent in rows); C ([1] lanceolatum; [2] oblongum; [3] ellipticus; [4] subulatus; [5] teres; [6] cylindricus; 7 bifurcates; [8] cuneatus; [9] angulato-ovatus; [10] semiteres); D ([1] trigonus; [2] teres; [3] compressus; [4] angulosus); E ([1] lanceolatum; [2] obtusum; [3] serratum; [4] filiformis; [5] repandus; [6] angulatus; [7] pinnatifidus; [8] ciliates) (Linnaeus, 1751)

2.1.4.1 Molecular taxonomic studies of *Sargassum* spp.

Sargassum C. Agardh is one of the morphologically most complex phaeophyceaen genera, and is divided into subgenera, sections, subsections, series and species group (Mattio *et al*., 2010). It is also a large genus with the largest member of taxa in Ochrophyta (Yoshida, 1983). The ‗genus' classification system is based on differences in macro-morphological characters and dated back to the 19th century. Those morphological characters may display variations within individual species and the taxonomic complexity of the genus is due to its highly polymorphic nature and phenotypic plasticity reported by several authors (Mattio & Payri, 2010). Taxonomic inconsistencies in their complex genera have caused confusion and presented many problems of classifications. Misidentification of certain species by morphological characters also occurs especially in young plants (Ajisaka *et al*., 1999). Thus, *Sargassum* species always create much hardship to taxonomists.

Phillips *et al*. (2005) had concluded that by using the more conservative region of the *rbc*LS operon (*rbc*L), they have successfully determined the closest lineage to the genus *Sargassum* while gaining other insights into generic to family level issues within the fucalean lineages and produced one of the first modern phylogenies for *Sargassum*.

Recently, DNA markers, in combination of nuclear (rDNA ITS-2), chloroplastic (partial rbcLS-operon), and mitochondrial (*cox3*), have been used for classification and analysis of phylogenetic relationships in *Sargassum* (Mattio *et al*., 2008; Mattio *et al*., 2009; Mattio *et al*., 2010). These studies resulted in several taxonomic changes, including the description of a new section which is *S*. sect. *Polycystae* Mattio & Payri (Mattio *et al*., 2009), abandoning subsections within *S*. sect. *Acanthocarpiceae*, synonymization of *S.* sect. *Acanthocarpicae* and sect. *Alacocarpicae* with sect.

Sargassum (Mattio *et al*., 2010), and thus managed to resolve the taxonomic incongruities of *Sargassum*.

2.2 Abiotic Stressors of Seaweeds in the Intertidal Zone

The shifting of environmental variables with the ongoing global climate change is expected to affect most of the organisms in marine coastal systems in terms of the performance, and the distribution (Wahl *et al*., 2011). Marine organisms living in these intertidal or shallow subtidal habitats are regularly exposing to recurring, harsh changes biotic and physical abiotic factors in the environment (Thomsen & Melzner, 2010). Biotic factors in benthic environment include biological interactions among seaweeds, between seaweeds and their epiphytes, of grazing from herbivores or human, whereas abiotic factors constitute the external physicochemical environment surrounding seaweeds, which include, changes in salinity, temperature, pH, wave action, irradiance etc (Wahl *et al*., 2011). These biotic and abiotic factors can affect the distribution of organisms in the intertidal zone. They can thus be considered as stressors. Seaweeds are usually immobile, thus they have to rely on cellular mechanisms to tolerate stress in order to survive in the intertidal zone (Dittami *et al*., 2009). Therefore, in order to clarify and evaluate the risks of global change, it is important to understand how organisms in harsh and fluctuating habitats cope with single and multiple stresses (Wahl *et al*., 2011).

Wave action is one of the most obvious stressors where waves exert mechanical stress on algae which grow in the intertidal zone. Thus, the only well adapted species, or no algae colonize very exposed areas of the shoreline at all. Seaweeds are susceptible to wave action for example tearing of laterals and/or detachment of whole thalli of seaweeds from substrate due to strong wave action brought by heavy rainfall which

directly inflict physical stress onto *Padina boryana* (Wichachucherd *et al*., 2009). However, thalli from exposed *F. vesiculosus* in both the Baltic Sea and the North Sea, were 30 % more resistant to tear and breakage as compared with conspecific from more sheltered sites (Wahl *et al*., 2011). Jonsson *et al*. (2006) reported that a negative correlation exists between thallus size in *Fucus vesiculosus* and *F. spiralis* with increasing water motion. Wave action strongly influences local distribution of *Fucus* spp. where it affect the gamete and zygote dispersal, the successsful attachment and survical of zygotes (Jonsson *et al*., 2006; Wahl *et al*., 2011).

Changing tide is also one of the important stressors. In the course of emersion, algae will start to desiccate. The effects of the desiccation are generally quite similar to those of hypersaline stress which normally happen during dry season. In both cases, the loss of the intracellular water will cause the increase in the intracellular ion concentration. Coralline bleaching is strongly induced by desiccation stress in the intertidal zones, which tighly coupled to high irradiance and high temperature stress at low tide (Martone *et al*., 2010).

Salinity is generally variable in the coastal region especially in the intertidal zone and most of the populations in the middle and uppermost intertidal are affected by salinity stress (Wahl *et al*., 2011). In response to saline/osmotic stress, it has been suggested that the sensors of the cell which probably located in the cell membrane can detect the changes. It is believed that instant responses of seaweeds to adverse environmental conditions involve excess production of reactive oxygen species (ROS) such as singlet oxygen ($^1O_2$), superoxide radical (O_2^-), hydroxyl radical (OH^-) and hydrogen peroxide (H_2O_2) (Pandithurai & Murugesan, 2014). Reactive oxygen species (ROS) are chemically reactive free radicals that play a crucial role in cell signaling and homeostasis, however, over production of ROS are toxic to cells and causes damage to

proteins, lipids, carbohydrates, DNA which eventually results in cell death (Gill & Tuteja, 2010; Ramakrishnan *et al*., 2016). Thus, rapid scavenging of ROS is essential in protecting cells from oxidative damage. It was suggested that saline stress (both hypo- and hypersaline) is one of the most important stresses in the intertidal zone (Collén *et al*., 2007). Rapid formation of ROS that reacts with most of the cellular components due to increased physiological stress conditions, thus need to be neutralized instantly. Changes in physiological process such as antioxidant enzymes (superoxide dismutase (SOD), glutathione reductase (GR), catalase (CAT) and ascorbate peroxidase (APX)) and non-enzymatic antioxidants (carotenoids, ascorbate and glutathione) were resulted in order to acclimatised to altered osmotic conditions (Liu & Pang, 2010). APX has been widely studied and has been identified directly involved in the scavenging in various organisms (Dabrowska *et al*., 2007; Caverzan *et al*., 2012; Urzica *et al*., 2012). APX is localized in the cytosol in *Euglena gracilis*, while chloroplast APX activity is restricted to thylakoids in bryophyte *Physcomitrella patens* and to the stroma in freshwater *Chlamydomonas reinhardtii* (Pitsch *et al*., 2010). SOD also is widely assumed to play a role in the detoxification of ROS which caused by enviromental stresses (Shiraya *et al*., 2015). Bowler *et al*. (1992) suggested that glutathione reductase may cooperate with SOD to remove superoxide radicals. Tammam *et al*. (2011) reported that chlorophytes *Dunaliella tertiolecta* had significant increase in SOD, CAT and GPX on exposure to extreme range of salinity which indicated that these enzymes were utilized for enzymatically scavenging of ROS. Thus, it is likely that SOD activity in *D. tertiolecta* increased for dismutation of superoxide ion and CAT and GPX together with ascorbate–glutathione cycle in extreme range of salinity degrade the subsequent generated H_2O_2. Portune *et al*. (2010) suggested that during growth of several rapidophytes, SOD and CAT play important protective roles against ROS.

Teo *et al*. (2009) investigated the transcriptomic changes of *Gracilaria changii* in response to hyper- and hypo-osmotic stresses using a cDNA microarray approach. Kumar *et al*. (2010) reported that almost two fold of increment in the contents of polyphenols, proline and the activities of antioxidative enzymes including APX, GR and SOD in both 6 days hypo- and hyper-salinity treatments of red alga *Gracilaria corticata*. The results suggested that the antioxidative enzymes, phycobiliproteins and mineral nutrients have potential role to combat the oxidative stress which induced by the salinity in *G. corticata*. Dittami *et al*. (2009) used EST-based microarray to study the transcriptome response of *Ectocarpus siliculosus* to three forms of abiotic stress: hyposaline, hypersaline, and oxidative stress. These studies lead a relatively partial and/ or complete picture of how plants perceive and respond to stress. On the other hand, heavy metal, especially copper stress, can also cause an intracellular accumulation of reactive oxygen species (ROSs).

While water temperature in the open ocean remains relatively constant throughout the day, water temperature in the absence of water during low tide or in the tidal pool is subjected to great variability (Dittami, 2010). As the temperature increase, the solubility of oxygen in water will decrease. This creating an environment deficient in oxygen and the pH in water might have some variability. The intensity and the availability of sunlight also have the influence in temperature.

Malaysia faces two monsoon wind seasons, the Northeast Monsoon (October to March) and the Southwest Monsoon (from late May to September) which the rainfall distribution pattern is determined by seasonal wind flow patterns. In 2004, Wong and Phang reported that the most important factor controlling the biomass production of *Sargassum* species in Cape Rachado, Port Dickson, Malaysia was the total rainfall.

2.3 RNA Extraction from Seaweeds

Isolation of high quality RNA from seaweed tissues which are rich in polysaccharides and polyphenols is always a difficult task. Phlorotannins, which are polyphenolic secondary metabolites, are present only in marine brown algae (Kamiya *et al*., 2010). They are stored in vesicles called physodes. The concentration of phlorotannin in brown seaweeds such as *Sargassum* spp. can vary with tissue type, age, species, seaweed size, abiotic factors such as water temperature, light intensity, nutrient levels and season (Steffy *et al*., 2013). The problem of isolating nucleic acid from seaweeds are the polysaccharides and secondary metabolites which are release after disruption of the cells (Ho *et al*., 1996). Moreover, composition of the seaweed cell wall showed a slight differences in the samples isolated from different organs or parts of the seaweed at different seasons in the year, and different species as well (Nyvall *et al*., 2003). In contrast to other organisms, extracting intact RNA from brown seaweed can be further complicated by enzymatic degradation by RNase. There have been numerous publications related to the difficulties of RNA extraction from seaweed and plant tissues which are rich in polysaccharides (Wong *et al*., 2007; Yao *et al*., 2009; Yockteng *et al*., 2013). These publications reported that even same species grown under different environments, or different species grown under the same environments, different conditions and/or different protocols were required for succesfull RNA extraction (Gehrig *et al*., 2000; Hong *et al*., 1997).

Several approaches were used to solve this problem including the use of cesium chloride density gradient ultracentrifugation (Mackay & Gallant, 1991). In addition, guanidinium thiocyanate, LiCl (lithium chloride) and CTAB (hexacetyltrimethyl ammonium bromide) were also added into the extraction buffers for RNA extraction from various seaweeds (Hong *et al*., 1995; Kim *et al*., 1997; Apt *et al*., 1995). However, RNA extraction from individual seaweeds remains a problem as majority of these

methods are genus-specific or species-specific. Wong *et al*. (2007) has optimized RNA extraction method for *S. binderi*, however, the same method failed to yield pure RNA for *S. polycystum*. Therefore, a new or improved method has to be developed or optimized to overcome the polysaccharide contaminations that are unique to individual species including *S. polycystum*.

2.4 Functional Genomics of Seaweeds

Functional genomics is a field of molecular biology, which focuses on the dynamic information such as gene transcription, translation, regulation of gene expression and protein-protein interactions. It describes the use of genome-wide assays to study gene and protein function, generally involving high-throughput methods rather than a more traditional "gene-by-gene" approach (Wikipedia, 2016a). A number of transcriptomic studies, which investigated the developmental process, physiological responses and establish the contribution of brown algae to diverse evolutionary lineages via secondary endosymbiosis event, are available on brown algae, including the diatom (Chan *et al*., 2011; Chan *et al*., 2012a), the genera *Ectocarpus* (Dittami *et al*., 2009), *Fucus* (Pearson *et al*., 2010), and *Saccharina* (Heinrich *et al*., 2012). On the other hand, the red algae including the genera *Gracilaria* (Teo *et al*., 2009), *Chondrus* (Collén *et al*., 2007), and *Porphyra* (Chan *et al*., 2012b) have contributed to trancriptomic studies.

Gene expression analyses at the mRNA level are one of the key components of functional genomics. A lot of methods were created to study the expression level of genes in seaweeds. These techniques includes expressed sequence tags (ESTs) (Shen *et al*., 2011; de Oliveira *et al*., 2012), next generation sequencing (Kim *et al*., 2014), microarray (Flöthe *et al*., 2014), quantitative reverse transcription-real time PCR (RT-qPCR), suppression substractive hybridization (SSH) (Shen *et al*., 2011), and other

analyses that enable the discovery of the gene of interest novel differentially expressed genes.

2.4.1 The Expressed Sequence Tag (EST) Approach and Its Applications to Seaweed Research

Since 1980, a lot of genome projects have been launched with the ultimate goal to identify all the genes of an organism. Human Genome Project (HGP) began officially in October 1990, followed by a number of ―model organisms", including bacterium *Escherichia coli*, the yeast *Saccharomyces cerevisiae* (Goffeau *et al*., 1996), the roundworm *Caenorhabditis elegans*, the fruitfly *Drosophila melanogaster* (Adams *et al*., 2000) and the mouse *Mus musculans* (Waterston *et al*., 2002). However, *S. cerevisiae* was the first eukaryotic genome that was completely sequenced (Goffeau *et al*., 1996). Till now, most of the Kingdoms have had their species completely mapped or being mapped. The data generated from genome sequencing become the fundamental resources for many molecular biology studies, which can be assessed via National Center for Biotechnology Information (NCBI). However, speed and cost are the major problems encountered by using whole genome sequencing method to complete the whole project. Therefore, an alternative method or a complementary method, expressed sequence tags (ESTs) sequencing, was proposed. The number of ESTs has increased dramatically in the past 15 years due to the faster and cheaper sequencing technologies.

In 1991, a high-throughput cDNA sequencing approach, was initiated in parallel to the ongoing human genome project for human gene discovery (Adams, 1991). The term ―expressed sequence tags" (ESTs) was at first time introduced to describe the strategy used. ESTs are unedited, randomly selected single-pass sequence reads and short sequences which consisted approximately 100-800 nucleotide base pair derieved from

cDNA libraries (Nagaraj *et al*., 2006). Messenger RNA (mRNA) sequences in the cell represent copies from expressed genes. RNA has to be reverse transcribed to double-stranded cDNA using reverse transcriptase, because RNA cannot be cloned directly. Thus, the resultant cDNA is cloned to make library, which represents a set of transcribed genes of cell, tissue or organism. Finally, the randomly picked cDNA clones from the library are subjected to single-pass sequencing with either end of the insert by using universal primers. Hence, ESTs are _tags' that represent partial sequences of coding sequences, only fragments of genes, which provide direct evidence for transcription profiles in cell, tissue or organism (Nagaraj *et al*., 2006; Harrow *et al*., 2009).

There are several public databases available for the deposition of EST data, and one of the largest EST sequence database is dbEST that was established by NCBI in 1992. At the time of writing (23 January 2016), 1,605 organisms with approximately 74,186,692 ESTs had been deposited in the dbEST (www.ncbi.nlm.nih.gov/dbEST/dbEST_summary/). A few algal genomes and EST projects were launched globally for the last few years. Majority of the data generated from these projects were available at the databases of the NCBI. In 2005, after being served by Sanger sequencing for over 30 years, the scientific community was introduced to next generation sequencing technologies, also called second generation sequencing (Gupta & Gupta, 2014). The advent of next-generation sequencing platforms have increased the speed and depth of transcriptome data dramatically at a significantly lower cost, especially for the majority of neglected organisms (Pareek *et al*., 2011).

For less-studied organisms, where whole genome sequencing is unavailable, the use of EST analysis is an attractive alternative. Currently, ESTs have been widely used on

many other organisms and become one of the most important resources for transcriptomic study. ESTs have diverse applications such as enabling gene discovery, understanding gene expression and regulation patterns of different tissues or developmental stages. It provides an initial important information of specific genes in an organism, where the specific genes may be expressing at critical or certain stages of their life cycle. Kenyon *et al*. (2003) have successfully identified a number of possible antigens through the generation of a few hundred ESTs from the sheep scab mite *Psoroptes ovis* and development of a few serodiagnostic test for sheep scab using recombinant protein Pso o 2, for the early detection of anti-*P. ovis* serum antibodies in sheep. The Pso o 2-based enzyme-linked immunosorbent assay (ELISA) was able to detect specific antibodies to *P. ovis* during experimental infestation prior to disease patency, indicating its utility for detecting sub-clinical infestation (Nunn *et al*., 2011). Characterization of stress genes which the ESTs generated from mangrove plant, *Acanthus ebracteatus* (Nguyen *et al*., 2006), may contribute to crop improvements in the future through the establishment of breeding program. The ESTs generated from fish, *Kryptolebias marmoratus*, have potential to be used as biomarkers for environment pollution and toxicity (Lee *et al*., 2007).

In the EST Database, the total volume of ESTs from higher plants is much larger compared with the total volume of ESTs from seaweeds (Ho *et al*., 2007; Gao *et al*., 2008; Narina *et al*., 2011; Sui *et al*., 2012). In 2012, red seaweeds *Porphyra umbilicalis* (L.) J. Agardh and *Porphyra purpurea* (Roth) C. Agardh provide comprehensive transcriptome data which comprising ca. 4.7 million ESTs reads (Chan *et al*., 2012b). Another 8088 ESTs from the red macroalga, *Gracilaria changii* were generated (Teo *et al*., 2007). As for green algae, 1898 ESTswere generated from the green macroalga *Ulva linza* (Stanley *et al*., 2005). In 2012, *de novo* assembly 382,884 reads from *U. linza* and generated 13,426 contigs (Zhang *et al*., 2012).

Genome data from a few red and green algal species are currently available. Genome data from red algae including the unicellular red alga *Cyanidioschyzon merolae* (Matsuzaki *et al*., 2004), unicellular red alga *Porphyridium purpureum* (Bhattacharya *et al*., 2013), the red seaweed, *Pyropia yezoensis* (Nakamura *et al*., 2013) and the red seaweed *Chondrus crispus* (Collén *et al*., 2013). As for green alga, *Chlamydomonas reinhardtii* (120 Mbp) is the only completed unicellular green alga nuclear genome (Merchant *et al*., 2007), whereas *Chlorella vulgaris* (Wakasugi *et al*., 1997) and *Helicosporidium* (Pombert & Keeling, 2010) have completed plastid genome and mitochondria genome, respectively. There is a general lack of genomic information available for phaeophyte seaweeds; the only published genomes available are of the filamentous alga, *Ectocarpus siliculosus* (Ectocarpales) (Cock *et al*., 2010), and of commercial important kelp, *Saccharina japonica* (Fucales) (Ye *et al*., 2015). While *E. siliqulosus* and *S. japonica* genomes are publicly available, the first ESTs for a brown alga was generated from *L. digitata* with the sequencing of 905 ESTs (Crépineau *et al*., 2000). In 2005, Roeder and co-workers generated additional 1985 ESTs from protoplasts of this alga (Roeder *et al*., 2005). This was followed by additional EST information derived from several brown algae: 1876 from *S. binderi* (Wong *et al*., 2007), 90,637 from *E. siliculosus* (Dittami *et al*., 2009), 12,000 from *F. serratus* and *F. vesiculosus* (Fucales, Ochrophyta) (Pearson *et al*., 2010) and 400,503 from *S. latissima* (Heinrich *et al*., 2012). However, the genomes of *E. siliculosus* and *S. japonica*, and the wealth of gene expression and function of these brown algae (Dittami *et al*., 2009; Cock *et al*., 2010; Ye *et al*., 2015) provide an excellent starting point for comparative analysis in other brown algae.

2.4.2 Real-time Quantitative Reverse Transcription Polymerase Chain Reaction (RT-qPCR) of Seaweed Genes

A wide number of high throughput techniques have been developed which allow the monitoring or quantification of expression levels of hundreds or thousands of genes simultaneously (O'Brien *et al*., 2012). Commonly, cDNA and oligonucleotide-based microarrays are used to measure transcripts at large scale. However, their sensitivity and accuracy are limited by low abundance transcripts and by the inability to distinguish between expressions of closely related genes (Costa *et al*., 2010). In contrast, real-time RT-PCR or quantitative RT-PCR (RT-qPCR) allows weakly expressed genes to be accurately quantified. It is the most sensitive, good reproducibility and specific method for assessing gene expression (Nolan *et al*., 2006). Quantitative real time polymerase chain reaction (RT-qPCR) is one of the most common approaches to functional genomics research, and it has become the most prevalent method applied to quantify assays of gene expression (Derveaux *et al*., 2010). However, a number of problems could limit its use, including amplification of unspecific products, primer-dimers, amplification efficiencies, hetero-duplex formation, etc (Pfaffl, 2004). One of the most important issues is the selection of suitable reference genes with stable expression, which is vital to effective normalization and the acquisition of accurate and meaningful biological data (Cinar *et al*., 2013; Kong *et al*., 2014). RNA may contain tissue enzyme inhibitors that result in reduced RT and PCR reaction efficiencies and generate unreliable and ―wrong" quantification results (Bustin, 2010).

RT-qPCR has been used to examine the expression level of mannuronan C5 epimerase gene expression in protoplasts after different times of cultivation and in sporophytes after different times of incubation in presence of homo-guluronates (oligoGG) elicitors (Tonon *et al*., 2008). Besides, this method is also used for data verification of gene expression analysis. In Ritter *et al*. (2014) study, 16 genes that

exhibited diverse expression patterns in the microarray analysis were analyzed by RT-qPCR. Eight of the genes were up-regulated, five down-regulated and three without significant changes during the copper stress. However, the confusion was solved through the verification of the gene expression of the 16 genes by using RT-qPCR. Only three genes coding for a putative zinc/iron permease (ZnFePer), a tyrosinase-like protein (TYR) and an expressed unknown protein, respectively, have different expression patterns obtained between microarray and RT-qPCR. Surprisingly, efforts to view the gene expression in *S. polycystum* have not been reported to date.

Normalization of a target gene expression is performed to offset technical limitations, such as RNA purity, poor efficiency of the synthesis of complementary DNA (cDNA), inaccurate quantification of RNA sample, and variability of pipetting. Even though normalization genes have been defined in many species, there is still no universal gene common to every organism (Kowalczyk *et al*., 2014). An ideal reference gene should show constant expression across all investigated samples regardless of cell type, developmental stage or other biological or experimental conditions (Goossens *et al*., 2005). However, a more robust, accurate and reliable approach to normalization is the use of multiple reference genes (Park *et al*., 2013; Zhu *et al*., 2013; Kowalczyk *et al*., 2014). Thus, the selection and validation of reference genes must be conducted before any meaningful RT-qPCR analyses can be performed. To this aim, a number of bioinformatics tools have been developed and widely used for choosing reference genes in various studies, which include GeNorm (Gimeno *et al*., 2014; Kowalczyk *et al*., 2014), BestKeeper (Pfaffl *et al*., 2004) and NormFinder (Andersen *et al*., 2004). GeNorm allows the most appropriate reference gene to be chosen by using the geometric mean of the expression of the candidate cDNA (Vandesompele *et al*., 2002) and is widely used to evaluate normalization genes (McMillan & Pereg, 2014). This software is freely available at http://www.genomebiology.com/2002/3/7/research/0s034

and the underlying principles are published by Vandesompele *et al*. (2002). Second program, BestKeeper is also select the least variable gene using the geometric mean but uses raw data instead of data converted to copy number (Pfaffl *et al*., 2004). It is available at http://www.genequantification.de/BestKeeper-1.zip. Norm-Finder, the third program is freely available on request. This program not only measures the variation but also ranks the potential reference genes by how much they differ between study groups, which is the extent by which they are affected by the experimental conditions. Defining this is essential as it can generate false results as discussed above.

CHAPTER 3: MATERIALS AND METHODS

3.1 Sample Collection and Preparation

Fresh plants of *Sargassum polycystum* C. Agardh were collected from coral reef at Teluk Kemang (N 02º26.422', E 101º51.355'), Port Dickson, Negeri Sembilan, Malaysia, on July and December, 2008 for expressed sequence tag (EST) generation purpose. Ten individual plant samples (mature plants without reproductive structures) were collected from each month. For quantification of gene expression purpose, four invidual plant samples (mature plants without reproductive structures) were harvested on 12th July, 2010 (rainy season) and 6th December, 2010 (dry season), respectively. These seaweeds transported back to the laboratory in a few hours upon collection, were maintained in natural seawater. Samples were washed carefully and packed into aluminium foil, and immediately snap-frozen in liquid nitrogen (N_2). All the samples were kept at -80 °C. In order to avoid the degradation of RNA, the collection, cleaning and storage of seaweed samples were accomplished on the same day. For each collection, the salinity and pH were measured *in situ* with a refractometer (Atago® MASTER-S/Millα, Tokyo, Japan), and a pH meter (CyberScan pH 600, Eutech Instruments, Thermo Fisher Scientific Inc., Ayer Rajah Crescent, Singapore), respectively. Meteorological data such as monthly rainfall was obtained from the Malaysian Meteorological Department (MMD), Petaling Jaya, Malaysia. Total rainfall for dry season (6th Dec) on 2010 was 0 mm (MMD, Appendix B), pH = 8.25 ± 0.04, salinity = 32 ppt, whereas total rainfall for rainy season (12th July) on 2010 was 1.8 mm (MMD, Appendix), pH = 8.45 ± 0.05, salinity = 31 ppt.

3.2 RNA Extraction Methods

In this study, five different RNA extraction methods were compared.

3.2.1 Method 1: Modified from Hong *et al.* (1997)

Ground frozen samples (2.5 g) were added to the extraction buffer (4 M guanidinium thiocyanate, 0.8 M LiCl, 10 mM ethylenediaminetetraacetic acid (EDTA), 2 % (*v/v*) β-mercaptoethanol, 0.6 % (*w/v*) *N*-lauroylsarcosine sodium (LSS), 0.2 % (*w/v*) PVPP) in a ratio of 1:10 [tissue (*w*)/buffer (*v*)]. The mixture was heated for 10 min at 55 °C, gently shaken for 1 h on ice, and spin at 10,000×*g* for 15 min at 4 °C. The RNA was precipitated with 0.1 volume 3 M sodium acetate (NaOAc) (pH 5.2) and 0.6 volume isopropanol at -20 °C overnight. Phenol/chloroform/isoamyl alcohol (PCI) (25:24:1, *v/v*, pH 8.0) extraction was performed, followed by precipitation using 2 volumes ethanol and 0.1 volume 3 M NaOAc (pH 5.2) at -20 °C overnight. The RNA was collected by centrifugation (10,000×*g*, 30 min, 4 °C). The RNA pellet was recovered by centrifugation and washed with ice-cold 70 % ethanol (*v/v*), air-dried and RNA pellet was dissolved in 50 µL DEPC-treated water.

3.2.2 Method 2: Modified from Kim *et al.* (1997)

Ground frozen tissue was added to the extraction buffer (1 % (*w/v*) SDS, 35 mM ethylene glycol tetraacetic acid (EGTA), 35 mM EDTA, 50 mM LiCl and 25 mM Tris-HCl (pH 8.0)) in a ratio of 1:10 (*w/v*). The mixture was mixed thoroughly and equal volume of PCI was added into the mixture and mixed. After centrifugation at 10,000×*g* at 4 °C for 15 min, the supernatant was made up to a final concentration of 4 M LiCl and left for 3 h at -20 °C. After centrifugation at 10,000×*g* for 30 min at 4 °C, the RNA was collected and then extracted with an equal volume of PCI twice. The aqueous phase was precipitated at -80 °C using 2.5 volumes ethanol and 0.1 volume 3 M NaOAc (pH 5.2) overnight. The RNA pellet was recovered by centrifugation at 10,000×*g* at 4 °C for

30 min, followed by washing with ice-cold 70 % ethanol and air-dried. Finally, the RNA pellet was dissolved in 150 µL DEPC-treated water.

3.2.3 Method 3: Modified from Wang *et al*. (2008)

Ground frozen tissue was added to the lysis buffer (2 M NaCl, 25 mM EDTA, 200 mM Tris (pH 8.0), 2 % (*v/v*) β-mercaptoethanol, 20 mM sodium borate, 2 % (*w/v*) polyvinylpolypyrrolidone (PVPP), 2 % (*w/v*) CTAB, 1 % (*w/v*) LSS) in a ratio of 1:10 (*w/v*). The mixture was vigorously mixed and incubated at 65 °C in a water bath shaker over 10 min with occasional shaking. One volume chloroform/isoamyl alcohol (CI) (24:1, *v/v*) was added into mixture, mixed, followed by centrifugation at 10,000×*g* for 15 min at 4 °C. The supernatant was transferred into another tube and 2 volumes of water-saturated PCI (25:24:1, *v/v*, pH 4.5) was added into the mixture and mixed. The mixture was centrifuged at 10,000×*g* at 4 °C for 10 min. The previously described step was repeated twice or until a clean interface was observed. Extraction with an equal volume of CI was performed and the aqueous phase was precipitated with 0.1 volume NaOAc (pH 5.2) and 2.5 volumes ethanol at -20 °C overnight. The RNA pellet was recovered by centrifugation, washed with ice-cold 70 % ethanol (*v/v*), air-dried and re-suspended in 100 µL DEPC-treated water.

3.2.4 Method 4: Wong et al. (2007)

Ground frozen tissue was added into the CTAB lysis buffer (100 mM Tris-HCl (pH 8.0), 2 % (*w/v*) CTAB, 20 mM EDTA, 1.5 M NaCl and 20 mM dithiothreitol (DTT) in a ratio of 1: 10 [tissue (*w*)/buffer (*v*)]. The mixture was mixed gently at 25 °C for 15 min prior to CI (24:1, *v/v*) extraction. The mixture was centrifuged at 10,000×*g* for 10 min.

The upper supernatant/aqueous phase was carefully transferred to a new tube without disturbing the cell-debris pellet. A total of 0.33 volume ethanol was added, subsequently by CI (24:1, *v/v*) extraction. β-mercaptoethanol and LiCl were added to a final concentration of 1 % (*v/v*) and 3 M, respectively and incubated at -20 °C for 3 h. The RNA pellet was obtained by centrifugation 10,000×*g*, 4 °C for 30 min and the resulting RNA pellet was dissolved in 400 µL RNase-free water. Extraction with an equal volume of PCI twice was performed and the aqueous phase was precipitated at -80 °C using 2 volumes ethanol and 0.1 volume 3 M NaOAc (pH 5.2) overnight. The RNA pellet was obtained by centrifugation at 10,000×*g*, 4 °C for 30 min. The RNA pellet was then washed with ice-cold 70 % ethanol, air-dried and was dissolved in 20 µL DEPC-treated water.

3.2.5 Method 5: Newly optimized method developed in this study

Finely ground frozen tissues were added into CTAB extraction buffer (100 mM Tris–HCl (pH 8.0), 2 M NaCl, 20 mM EDTA, and 2 % (*w/v*) CTAB) and 50 mM DTT was added prior to extraction in a ratio of 1:10 [tissue (*w*)/buffer (*v*)]. The mixture was mixed vigorously prior to CI extraction. The mixture was centrifuged for 10 min at 10,000×*g*. The supernatant was transferred to a new tube and 0.3 volume of absolute ethanol was gently added, followed by CI extraction. The mixture was then centrifuged at 10,000×*g* for 10 min. The supernatant was transferred to a new tube and LiCl was added to a final concentration of 2 M, before incubation at −80 °C for 2 h. The RNA pellet obtained by centrifugation at 10,000×*g* for 30 min at 4 °C was then dissolved in 400 µL RNase-free water. CI extraction was repeated once, followed by the addition of LiCl to a final concentration of 2 M and precipitated at -80 °C for 2 h. The RNA pellet

was recovered by centrifugation and then washed with ice-cold 70 % ethanol and air-dried. Finally, the RNA pellet was dissolved in 20–50 μL DEPC-treated water.

3.3 Quantification of Total RNA

3.3.1 Spectrophotometer

The RNA was quantified using a NanoPhotometer™ UV/Vis spectrophotometer (Implen GmbH, Schatzbogen, Germany). The RNA purity and concentration were assessed by determining the spectrophometric absorbance of samples at 230, 260, and 280 nm, ratio of A_{260}/A_{280}, and ratio of A_{260}/A_{230} (Sim *et al*., 2013). One unit of absorption at 260 nm represents 40 μg/mL of RNA. The ratio of A_{260}/A_{280} provides an indication of protein contamination and the ratio of A_{260}/A_{230} indicates polysaccharide/ polyphenolic contamination (Chan *et al*., 2004; Japelaghi *et al*., 2011).

3.3.2 Formaldehyde-Agarose Gel Electrophoresis

Formaldehyde-agarose gel (1.2 %) was prepared by melting agarose in 1x F buffer (20 mM MOPS pH 7.0, 1 mM EDTA, 5 mM NaOAc). Formaldehyde was then added to the agarose solution to a final concentration of 6 % (*v/v*) after the solution cooled down to about 55 °C. The warm agarose solution was then poured into the mould. A total of 1-2 μg RNA was added to the sample buffer (6 % *v/v* formaldehyde, 50 % *v/v* formamide and 1x F buffer). The mixture was heated for 10 min at 65 °C, and loaded into the gel after 0.5 μL ethidium bromide (10 mg/mL) and 1 μL of saturated bromophenol blue were added. The electrophoresis was carried out in running buffer (1x F buffer and 6 % formaldehyde) at 5 V/cm.

3.4 mRNA Isolation

In this study, μMACS^{TM} mRNA Isolation Kit (Miltenyi Biotec, Gladbach, Germany) was used to isolate mRNA from the total RNA. Elution buffer was pre-heated to 65 °C using a heating block, and lysis/binding and wash buffer were warmed up to room temperature. The total RNA was heated to 65 °C for 5 min and then chilled briefly on ice, before it was diluted with 1 volume lysis/binding buffer. The final volume was between 250 – 1000 μL. About 200 μg or less total RNA was mixed with 50 μL oligo (dT) microbeads by pipetting up and down 2-3 times or mixed shortly. A MACS column type μ was placed in the magnetic field of the μMACS separator. The column was prepared by rinsing with 100 μL lysis/binding buffer. After that, the total RNA was applied to the column matrix. The column was rinsed with 200 μL lysis/binding buffer to remove proteins and DNA, followed by 4x 100 μL wash buffer to remove rRNA and DNA. For the elution of mRNA, 120 μL of 65 °C pre-heated elution buffer was applied. The mRNA eluted from the column was collected into a 1.5 mL tube and then precipitated in 0.1 volume 3 M NaOAc (pH 5.2) and 2.5 volume ethanol. The mRNA was stored at -80 °C.

3.5 Quantification of mRNA

A serial dilution of RNA ladder (1 μg/μL) was performed to prepare a few different concentrations of RNA. Each diluted RNA ladder (1 μL) was dotted on the petri dish, containing 0.8 % (*w/v*) of solid agarose gel with ethidium bromide (1 μg/mL). Then, 1 μL of mRNA sample was dotted onto the agarose gel and air-dried. The plate was viewed under UV light to estimate the concentration of the mRNA samples by comparing the intensities of the standard RNA of different concentrations that have

been prepared earlier. Besides, quantification of mRNA was also determined by using a NanoPhotometer^{TM} UV/Vis spectrophotometer (Implen, Schatzbogen, Germany).

3.6 cDNA Library Construction and EST Generation

The construction of cDNA library of *S. polycystum* was conducted using the ZAP-cDNA® Gigapack® III Gold Cloning Kit (Stratagene, La Jolla, CA, USA) with some modifications to the instructions provided by the manufacturer.

3.6.1 First-strand cDNA Synthesis

In an RNase-free microcentrifuge tube, the following reagents were added in order: 7.5 µL of poly(A)$^{+}$ RNA (5.6 µg), 3 µL of first-strand methyl nucleotide mixture (10 mM dATP, dGTP and dTTP plus 5 mM 5-methyl dCTP), 2 µL of oligo (dT) linker-primer (1.4 µg/µL), which was made up to 31.5 µL with DEPC-treated water. The reaction was mixed gently, heated to 65 °C for 5 min and quickly put on ice for 2 min. The contents were briefly centrifuged. Ten µLof 10x first-strand buffer, 5 µL 0.1 M DTT and 1 µL RNase inhibitor (40 U/µL) were added into the mixture and mixed gently, followed by incubation at 42 °C for 2 min. 2.5 µL of Superscript^{TM} RNase H^{-} reverse transcriptase (200 U/µL) (Invitrogen, Carlsbag, CA, USA) was added to the first-strand synthesis reaction and mixed gently by pipette. The mixture was incubated at 42 °C for 50 min, followed by 70 °C for 15 min to denature the enzyme, and then on ice. Two µL of the first-strand synthesis reaction was stored at -20 °C for gel electrophoresis later.

3.6.2 Second-strand cDNA Synthesis

A mixture of 48 µL of first-strand cDNA synthesis reaction, 20 µL of 10x second-strand buffer, 6 µL of second-strand dNTP mixture (10 mM dATP, dGTP, dTTP and 26 mM dCTP), 116 µL sterile distilled water, 2 µL RNase H (1.5 U/µL) and 11 µL of DNA polymerase I (9.0 U/µL) were prepared. The mixture was mixed gently, centrifuged, and incubated for 2.5 h at 16 °C. The reaction tube was then placed immediately on ice.

3.6.3 Blunting the cDNA Termini

Second-strand synthesis reaction was mixed with 23 µL of blunting dNTP mixture (2.5 mM dATP, dGTP, dTTP and dCTP) and 2 µL of cloned *Pfu* DNA polymerase (2.5 U/µL). The reaction was mixed immediately and centrifuged in a microcentrifuge. The reaction was incubated for 30 min at 72 °C. This was followed by the addition of 200 µL of phenol-chloroform (1:1) (pH 7-8) and the mixture was mixed, and centrifuged at maximum speed at room temperature for 2 min. The upper layer was transferred to a new tube and 20 µL of 3 M NaOAc and 400 µL of 99 % ethanol were added to the solution. The mixture was mixed and incubated at -20 °C for overnight.

The DNA pellet was centrifuged at 12,000×*g* for 60 min at 4 °C. The supernatant was discarded and washed with 500 µL 70 % ethanol. The ethanol was aspirated after centrifugation at maximum speed for 2 min at room temperature. The pellet was then resuspended in 9 µL of *Eco*R I adapters (0.4 µg/µL) and incubated at 4 °C for 45 min to dissolve the cDNA pellet, followed by ligation to the *Eco*R I adapter. One microlitre of second-strand cDNA synthesis reaction was transferred to a separate tube and kept at -20 °C for further use.

3.6.4 Ligation of *Eco*R I Adapters

The following components, 1 µL of 10x ligase buffer, 1 µL of 10 mM rATP, and 1 µL of T4 DNA ligase (4 U/µL) were added to the tube containing the blunted cDNA and the *Eco*R I adapters. The reaction was centrifuged and incubated at 4 °C for 2 days. The enzyme activity was inactivated at 70 °C for 30 min, followed by phosphorylation of the *Eco*R I ends.

3.6.5 Phosphorylation of the *Eco*R I Ends

After the ligase was heat inactivated, this 11 µL cDNA ligated to *Eco*R I adapter mixture was centrifuged in a microcentrifuge for 2 seconds. The reaction was cooled at room temperature for 5 min. The adapter ends was phosphorylated by adding the following components: 1 µL of 10x ligase buffer, 2 µL of 10 mM rATP, 5 µL of sterile water and 2 µL of T4 polynucleotide kinase (5 U/µL). The mixture was incubated for 30 mins at 37 °C, and followed by heat-inactivated kinase for 30 min at 70 °C. The mixture was centrifuged for 2 seconds and equilibrated to room temperature for 5 min.

3.6.6 Digestion with *Xho* I

A reaction of 21 µL cDNA with phosphorylated *Eco*R I ends, 28 µL of *Xho* I buffer supplement and 3 µL of *Xho* I (40 U/µL) were incubatedfor 1.5 h at 37 °C. The cDNA was precipitated by 0.1 volume NaOAc (pH 5.2) and 2.5 volume of 99 % ethanol. The mixture was incubated at -20 °C overnight.

The DNA pellet was centrifuged at a maximum speed 12,000×*g* for 60 min at 4 °C. The supernatant was discarded and washed by 500 µL 70 % ethanol. The air-dried pellet was dissolved in 20 µL sterile distilled water.

3.6.7 Size Fractionation by Agarose Gel Electrophoresis

A total of 20 µL cDNA (from Section 3.6.6) was electrophoresed in 1 % agarose gel, with λ-*Pst* I as molecular size marker. QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany) was used to extract and purified the cDNA fraction (500 bp to 10 kbp) from the gel. The gel was excised with a scalpel, transferred to a 50 mL tube and the weight of gel slice was recorded. Buffer QG was added to the gel in 1:3 (*w/v*). The tube was incubated at 50 °C for 10 min or until the gel was completely dissolved.

A maximum volume of 800 µL of the solution was transferred to a column with a 2 mL collection tube. The column was centrifuged at 13,000×*g* for 1 min at room temperature. The flow-through in the 2 mL-collection tube was discarded. The same procedure was repeated until all the solution flowed through the column. The column was added with 750 µL buffer PE and incubated for 5 min. The column was centrifuged at 13,000×*g* for 1 min at room temperature. The collection tube was replaced by a new 1.5 mL centrifuge tube, followed by the addition of 200 µL of TE buffer and incubated for 1 min at room temperature. The cDNA was eluted by centrifugation at 13,000×*g* for 1 min at room temperature. The fractionated cDNA was precipitated overnight at -20 °C by adding 0.1 volume NaOAc (pH 5.2) and 2.5 volume 100 % ethanol.

The DNA pellet was centrifuged at 13,000×*g* for 1 h at 4 °C. The supernatant was discarded, the pellet was washed with 500 µL 70 % ethanol, and centrifuged again at

12,000×*g* for 2 min at 4 °C. The air-dried cDNA pellet was dissolved in 3 µL sterile distilled water.

3.6.8 Ligation of DNA to the Uni-Zap XR Vector

Approximately 100 ng of fractionated cDNA (Section 3.6.7), 0.5 µL 10x ligase buffer, 0.5 µL 10 mM rATP (pH 7.5), 1.0 µL Uni-ZAP XR vector (1 µg/µL) (Appendix C, D), sterile water (to a final volume of 4.5 µL) and 0.5 µL T4 DNA ligase (4 U/µL), were mixed gently, and followed by centrifuged briefly. The reaction was incubated at 4 °C for 2 days.

3.6.9 Preparation of the Host Bacteria

The *Escherichia coli* XL1-Blue MRF' cells (Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*) *173endA1supE44thi-1 recA1 gyrA96 relA1 lac* [F' *proAB lacl*^{q}*ZΔM15* Tn*10* (Tet^{r})] glycerol stock was streaked ona tetracycline agar plate (15 µg/mL) and incubated at 37 °C for overnight. A single colony was picked and inoculated in 20 mL LB broth with 15 µg/mL of tetracycline. The culture was grown at 37 °C, shaken at 200 rpm overnight. A mixture of 2 mL overnight culture, 18 mL LB broth supplemented with 10 mM $MgSO_4$ and 0.2 % (*w/v*) maltose and 15 µg/mL tetracycline, were incubated at 37 °C with shaking at 200 rpm for 3-4 h (OD_{600} less 1.0). The cells were centrifuged at 6000 rpm at 4 °C for 10 min. The cell pellet was then resuspended in 10 mM $MgSO_4$ until the OD_{600} was adjusted to 0.5. The same method was repeated for the preparation of *E. coli*SOLR cells (e14^{-}(McrA^{-}) Δ(*mcrCB-hsdSMR-mrr*) *171 sbcC recB recJ uvrC umuC*::Tn*5* (Kan^{r}) *lac gyrA96 relA1 thi-1 recA1 thi-1 endA1* λ^{R} [F' *proABlacl*^{q}*ZΔM15*] Su^{-} (nonsuppressing). However, 50 µg/mL of kanamycin instead of tetracycline was used.

The purpose of using antibiotic in the agar is allows for the selection of only those bacteria with the specific antibiotic resistance which is usually conferred by a plasmid carrying the antibiotic resistance gene (Addgene, 2016).

3.6.10 DNA Packaging

ZAP-cDNA® Gigapack® III Gold Cloning Kit (Stratagene, La Jolla, Ca, USA) was used to pack the cDNA ligated to the Uni-ZAP vector. The packaging extract vial was removed from -80 °C. The ligation mixture (Section 3.6.8) was added immediately to the packaging extract when it began to thaw. The mixture was mixed gently by pippeting without introducing bubbles. The reaction was centrifuged for 5 seconds, and incubated at 22 °C for 2 h. A total of 500 µL SM buffer (100 mM NaCl, 8 mM $MgSO_4 \bullet 7H_2O$, 50 mM Tris-Cl (pH 7.5), 0.01 % gelatin solution), was added and followed by addition of 20 µL chloroform. The reaction was mixed gently, centrifuged and kept at 4 °C.

3.6.11 Plating and Titering

For primary library titering, a serial dilution of the phage was prepared. Serial dilutions: 1x, 10x, 100x and 1000x was prepared. The lambda phage was diluted with SM buffer. From each dilution, 1 µL of lambda phage was added into 200 µL of *E. coli* XL1-Blue MRF' cells (OD_{600} = 0.5) and incubated for 15 min at 37 °C. The mixture was added into 15 mL-tube, which contained 9 mL melted NZY top agar (85 mM NaCl, 8 mM $MgSO_4 \bullet 7H_2O$, 0.5 % (*w/v*) yeast extract, 1 % (*w/v*) NZ amine (casein hydrosylate), 0.7 % (*w/v*) agarose) at 48 °C. The top agar mixture was plated on NZY agar plates containing 2.5 mM IPTG and 5 mg/mL X-gal immediately. The solidified

agar plates were inverted and incubated at 37 °C for 6 to 8 h. After the incubation period, the plaques were counted.

3.6.12 Amplification of the Uni-ZAP XR cDNA Library

The host cell, *E. coli* XL1-Blue MRF' (Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*) *173endA1supE44thi-1 recA1 gyrA96 relA1 lac* [F' *proAB lacI*^{q}*ZΔM15* Tn*10* (Tet^{r})] was prepared one day before the experiment. A total of 20 tubes and 150 mm diameter NZY agar plates were prepared. For each 150 mm diameter NZY agar plate, 600 µL cells (OD_{600} = 0.5) and 5 x 10^4 pfu of bacteriophage (24.5 µL) from the cDNA library were needed. The mixture was incubated for 15 min at 37 °C to allow attachment of phage to host cell, followed by adding the mixture into a tube containing 6.5 mL of NZY top agar at 48 °C. The mixture of the top agarose was spread evenly on the 150 mm diameter NZY agar plate. This method was repeated for the other NZY agar plates. The plates were incubated at 37 °C for 6-8 h to prevent the plaques from growing more than 2 mm in diameter touching each other. After the plaques formed, the plates were then overlaid with 8 mL SM buffer each. The plates were then stored overnight at 4 °C to allow the phages to diffuse into the SM buffer.

On the following day, the plates were put on a rocking platform for 1 h. The bacteriophage suspension was recovered from each plate and pooled into a sterile Falcon 50 mL conical centrifuge tube. Each plate was then rinsed with SM buffer and transferred to the same tube. The debris was removed by adding 5 % (*v/v*) chloroform, mixed well and incubated at room temperature for 15 min. The tube was then centrifuged for 10 min at 500×*g*. The supernatant was recovered and transferred to a new tube. For long term storage, 7 % (*v/v*) DMSO was added and stored at -80 °C. For short term storage, chloroform was added to give 0.3 % (*v/v*) final concentration and

stored at 4 °C. The titer of amplified library was checked using host cells and a serial dilution of the cDNA library.

3.6.13 Single Clone *In Vivo* Excision

By using ExAssist helper phage with *E. coli*SOLR strain (e14$^-$(McrA$^-$) Δ(*mcrCB-hsdSMR-mrr*) *171 sbcC recB recJ uvrC umuC::*Tn*5* (Kan^r) *lac gyrA96 relA1 thi-1 recA1 thi-1 endA1* λ^R [F' *proABlacI*^{q}Z*ΔM15*] Su$^-$ (nonsuppressing), the pBluescript phagemid was excised from Uni-ZAP XR vector. Firstly, 60 µL of *E. coli* XL1-Blue MRF' cells (Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*) *173endA1supE44thi-1 recA1 gyrA96 relA1 lac* [F' *proAB lacI*^{q}Z*ΔM15* Tn*10* (Tet^r)] (OD_{600} = 1.0), 63 µL of phage stock and 1 µL ExAssist helper phage were mixed and incubated at 37 °C. Then, 750 µL LB broth was added and incubated for 3-4 h at 37 °C with shaking at 200 rpm. The mixture was heated at 70 °C for 20 min, followed by centrifugation at 1000×*g* for 15 min at room temperature. The phage supernatant was transferred to a new sterile tube. A volume of 50 µL *E. coli* SOLR cells (OD_{600} = 1.0) and 25 µL phage supernatant were added into a tube and incubated at 37 °C for 15 min. Fifteen microlitre of the mixture was plated on LB-ampicilin agar plate (50 µg/mL) and incubated at 37 °C overnight.

After *in vivo* excision, a few colonies were randomly picked from each plate which had been supplemented with 50 µg/mL ampicilin, using a sterile toothpick and sub-cultured in 5 ml LB broth supplemented with 50µg/mL ampicilin. The LB cultures were incubated in a shaking incubator at 37 °C for 12 to 16 h. The cultures were transferred into a 1.5 mL microfuge tube. The cells were centrifuged at 9000×*g* for 3 min at room temperature.

3.6.14 Mass Excision

E. coli XL1-Blue MRF' and *E. coli* SOLR cells with OD_{600} = 1.0 (8 x 10^8 cells/mL) were prepared one day before mass excision. A mixture of 4.8 x 10^{10} pfu of amplified lambda bacteriophage library, 1 µL of XL1-Blue MRF' cells (8 x 10^8 cells/mL) and 1 µL of ExAssist helper phage (2 x 10^9 pfu) in a ratio 1:10:100 was prepared in a 50 mL-tube. The mixture was incubated at 37 °C for 15 min. Then, 20 mL LB broth was added and incubated at 37 °C with shaking for 3 h. The activity of the ExAssist helper phage was heat-inactivated at 70 °C for 20 min. The debris was centrifuged at 6000×*g* for 10 min, followed by transferring the supernatant to a new sterile 1.5 mL tube with the addition of 7 % Dimethyl Sulfoxide (DMSO) and kept at -20 °C. Serial dilutions of 1x, 100x, 1,000x and 10,000x of excised phagemids were prepared. From each dilution, 1 µL excised phagemids were mixed with 200 µL SOLR cells (OD_{600} = 1.0) in 10 mM $MgSO_4$, followed by incubation at 37 °C for 15 min. After that, 100 µL of the cell mixture was plated on LB-ampicilin plates (50 µg/mL) and incubated at 37 °C for 16 h.

3.6.15 Plasmid Isolation

Plasmid DNA-spin™ purification kit (iNtRON, Inc., Seongnam-si, Korea) was used for plasmid isolation. The kit provided all the chemical reagents. Bacterial colonies were randomly picked from the overnight culture plate and grown in 15 ml culture tube which contained 5 mL LB broth with ampicilin (50 µg/mL), and incubated overnight with shaking. The bacterial culture was harvested by centrifugation at 13,000×*g* for 30 seconds at room temperature and the supernatant was discarded. Then, the pellet was re-suspended in 250µL of resuspension buffer, mixed cell until no clumps remain. This was followed by the addition of 250 µL lysis buffer and the mixture was mixed gently. After that, 350 µL of neutralization buffer was added and mixed by inverting the tube

several times, and centrifuged at 13,000×*g* for 10 min at 4 °C. After centrifugation, the supernatant was transferred to a spin column with the collection tube. The spin column with the supernatant was centrifuged at 13,000×*g* for 1 min. Filtrate in the collection tube was discarded and the spin column was placed in the same collection tube. An additional 700 μL of Washing Buffer B was added and centrifuged at 13,000×*g* for 1 min. Filtrate in the collection tube was discarded and the spin column was placed in the same collection tube. The spin column was centrifuged at 13,000×*g* for 1 min to dry the filter membrane. The collection tube was discarded and the spin column was placed in a new 1.5 mL microfuge tube. Elution buffer (50 μL) was added to the upper reservoir of the spin column, and incubated for 1 min. Then, the tube was centrifuged at 13,000×*g* for 1 min. The elution was kept at -20 °C.

3.7 Verification of cDNA Inserts

3.7.1 Polymerase Chain Reaction (PCR)

The PCR reaction [1x DyNAzyme™ buffer, 200 μM dNTP, 1 U DyNAzyme™ II DNA polymerase (Finzymes, Espoo, Finland), 200 μM T3 primer, 200 μM T7 primer and autoclaved distilled water] was prepared. PCR was performed at 94 °C for 2 min, 35 cycles of 94 °C for 1 min, 55 °C for 1 min, 72 °C for 2 min, followed by 72 °C for 7 min. A negative control without template was included with PCR amplifications. The PCR products were kept at -20 °C for further analysis. 1 % (*w/v*) agarose gel electrophoresis was carried out to analyze the PCR products, with 1kb DNA ladder (Fermentas, Ontario, Canada).

3.8 Sequencing and transcripts assembly

Sequencing of plasmid DNA was performed by using sequencing services provided by Solgent, Co., Ltd., Daejeon, Korea. A total of 2592 plasmids have been extracted, and 2586 plasmids were subjected to automated cycle sequencing. Automated cycle sequencing was performed by using T3 universal primer and ABI PRISM® BigDye™ Terminator version 3.1 (Applied Biosystem, California, USA), and electrophoresed on DNA Sequencer Model ABI 3730XL or ABI PRISM 377 (Applied Biosystem, California, USA). Two computer files were generated from each sequencing run, that is a chromatogram file and a plain text file. Raw chromatogram sequences were first individually processed using Phred software (CodonCode Aligner/ Fourier transform) which was linked to base calling, read trace data, and assignment of quality values to bases. Trimming of low quality bases, vectors and potential polyA regions was performed using FastX Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Sequences < 100 bp after trimming were eliminated from subsequent analysis. To reduce redundancy in transcripts coding for similar/identical functions, these EST sequences were assembled into contigs using CAP3 (Huang & Madan, 1999) at default setting. ESTs with an overlap of > 50 bp over 100 bp were assembled into contiguous sequences _contig', and the non-overlapping sequences were treated as _singletons'.

3.9 Functional Annotation

To identify functional homologs of the assembled ESTs (unigenes) from *S. polycystum*, each unigene (contig and singleton, respectively) was compared against the Swiss-Prot protein database (http://www.uniprot.org/) and the NCBI non-redundant (nr) protein database (http://blast.ncbi.nlm.nih.gov/Blast.cgi), using Blast2GO (Conesa & Götz, 2008). Putative function of each unigene was annotated based on its significant

hits (BLASTX, $E \leq 10^{-5}$); potential bacterial, diatoms and oomycetes contaminants were removed from the final unigenes set.

Mapping of the sequences according to Gene Ontology (GO) was conducted using Blast2GO (Conesa & Götz, 2008). GO mapping was performed by using the BLASTX hits derived from NCBI nr proetin database. Blast2GO was used to identify protein families, domains, regions and sites by InterProScan (Quevillon *et al*., 2005), which associated GO terms were merged with those identified by alignment to NCBI nr protein database. Enzyme codes were assigned to unigenes by Blast2GO based on associated GO terms providing results for Enzyme Commision (EC) numbers and their KEGG pathways. Enzyme commision (EC) number was mapped to proteins at e-value-hit-filter = $1.0e^{-6}$, annotation cut-off = 55, GO weight = 5, and HSP-Hit coverage cut-off = 40.

3.10 Sequence Analysis

The full transcript sequence encoding putative GDP mannose 4,6 dehyratase (GM46D) (Contig159), full transcript sequence encoding putative GDP-4-keto-6-D-mannose-epimerase-4-reductase (GFS) (Contig87), partial transcript sequence encoding mannuronan C5-epimerase (MC5E-1) (SP01411), and partial transcript sequence encoding putative mannuronan C5-epimerase (MC5E-2) (SP02271) were analysed using Bioedit Sequence Alignment Editor version 7.0.9, respectively (Hall, 1999). Amino acid composition and molecular weight determination were performed using Bioedit v7.0.9 (Hall, 1999). Proteins of Stramenopiles that are homologous to the proteins coded by MC5E-1, MC5E-2, GM46D-Contig159 and GFS-Contig87 were inferred based on BLASTP ($E < 10^{-10}$) search against NCBI nr database. Amino acid sequences of GM46D, GFS and MC5Es from other organisms were downloaded from

the National Center for Biotechnology Information (NCBI) and UnitProt (www.uniprot.org), respectively. Multiple sequence alignment for each of these homologous set were performed using ClustalW (Thompson *et al*., 2002), respectively. Partial protein sequences of MC5Es from *Sargassum polycystum* were aligned with ClustalW algorithm using amino acid sequences from *Ectocarpus siliculosus* (brown seaweed; CBJ32776; CBN77785), *Laminaria digitata* (brown seaweed; CAD42950; CAD42945), *Saccharina japonica* (brown seaweed; BAF80877). As for protein sequence of GM46D (Contig159) of *S. polycystum*, protein sequence was aligned with ClustalW algorithm using amino acid sequences from *E. siliculosus* (brown seaweed; D8LKT9), *Aureococcus anophagefferens* (marine picoplankton; F0Y4G6), *Thalassiosira pseudonana* (diatom; B8C1P6), *Blastocystis* sp. (unicellular protist; XP_014525236.1). Protein sequence of GFS (Contig87) from *S. polycystum* was aligned with ClustalW algorithm using amino acid sequences from *E. siliculosus* (brown macroalga; D7FW27), *A. anophagefferens* (marine picoplankton; XP_00903743), *N. gaditana* (marine microalga; EWM25919.1), *L. digitata* (brown macroalga; Q9SMD7).

The phylogenetic trees analyses of the GM46D and GFS were conducted, respectively. Representative amino acid sequences of GM46D and GFS were compiled from UniProt, GenBank protein database and our data from *S. polycystum*. Sequences were aligned using default setting MEGA 7 (www.megasoftware.net), with subsequent manually adjusted. The phylogenetic relationships of the amino acid sequences were constructed using MrBayes version 3.2.6 (Ronquist *et al*., 2011) with runs of Markov chain Monte Carlo (MCMC) chains, for two million generations, with sampling frequency of 100 and diagnostic frequency of 1000 generations. Convergence was assessed manually with the standard deviation of split frequencies below 0.01.

Protein domain of GM46D (Contig159), GFS (Contig87) and MC5Es (SP01411 and SP02271) were identified using PROSITE database (http://www.expasy.ch/tools/scanprosite/) (de Castro *et al*., 2006). In addition, subcellular localization of protein GM46D and GFS from *S. polycystum* were predicted using WoLFPSORT (http://wolfpsort.org/) (Horton *et al*., 2007) and HECTAR (Gschloessl *et al*., 2008), respectively. Signal peptide prediction for each protein sequence (SP01411, SP02271, Contig159 and Contig87, respectively) was analysed using SignalP 4.0 (Petersen *et al*., 2011).

3.11 cDNA Synthesis for RT-qPCR (Quantitative Reverse Transcriptase-Polymerase Chain Reaction)

Total RNA samples were treated with DNase I (New England BioLabs, Ontario, Canada) according to the recommendations of the manufacturer with some modifications. Thirty microgram of total RNA was resuspended in 1x DNase I Reaction buffer (10 mM Tris-HCl, 2.5 mM $MgCl_2$, 0.5 mM $CaCl_2$, pH 7.6 at 25 °C) to a final volume of 100 µL. One unit of DNase I was added, mixed thoroughly and incubated at 37 °C for 10 min. Five mM final concentration of EDTA was added. Phenol/chloroform/isoamyl alcohol (PCI) (25:24:1, *v/v*, pH 8.0) extraction was performed, followed by precipitating using 2.5 volumes ethanol and 0.1 volume 3 M NaOAc (pH 5.2) at -20 °C overnight. The DNase I-treated RNA pellet was recovered by centrifugation at 10,000×*g* at 4 °C for 30 min, and followed by washing with ice-cold 70 % (*v/v*) ethanol and air-dried. The RNA pellet was dissolved in 50 µL DEPC-treated water.

DNase I-treated total RNA samples (3 µg each) were reverse-transcribed using AffinityScript QPCR cDNA Synthesis Kit (Strategene, USA) according to the recommendations of the manufacturer.

First-strand cDNA synthesis reaction was prepared in a microcentrifuge tube by adding the following components in order: RNase-free H_2O to a total volume of 20 µL, 1x of first strand cDNA synthesis master mix, 0.3 µg oligo(dT) primer, 1.0 µL of AffinityScript RT/RNase Block enzyme mixture, and 3 µg of total RNA. The reaction was incubated at 25 °C for 5 min to allow primer annealing. Then, the reaction was incubated at 42 °C for 15 min to allow cDNA synthesis, followed by incubation at 95 °C for 5 min to terminate the cDNA synthesis reaction. The completed first-strand cDNA synthesis reactions were placed on ice for immediate use in qPCR. For long-term storage, the reactions were kept at -20 °C.

Prior to use in RT-qPCR, the first strand cDNA was quantified using NanoPhotometer (IMPLEN GmBH, USA), diluted to a concentration of 200 ng μL^{-1} with nuclease-free water immediately and stored at -20 ºC until used.

3.12 Selection of Endogenous Genes

In order to normalize the real-time RT-qPCR results accurately, the expression stability of four endogenous genes (alpha tubulin (AT), ribosomal protein L3a (RPL3a), ribosomal protein S15a (RPS15a) and translation elongation factor 1 alpha (TEF1A)) in *S. polycystum* samples at different seasons (rainy and dry seasons, respectively) were assessed using geNorm version 3.5 software (Vandesompele *et al*., 2002). Gene expression stability was evaluated using the geNorm software following the procedures of Vandesompele *et al*. (2002). Briefly, stability (M = gene-stability measure) refers to

the constancy of the expression ratio between 2 or more housekeeping genes among all samples tested. The more stable the expression ratio among 2 or more genes, the more likely that the genes are appropriate internal controls. The lower the M value, the higher the stability. For each gene, a pair of oligonucleotide sequences were designed to be annealing at the 3' UTR of the genes when the sequence was known, or to be annealing at the 3' coding sequence using Primer3 software (Rozen *et al*., 2000).

RT-qPCR was carried out by using the Brilliant III Ultra-Fast SYBR® Green QPCR Master Mix (Agilent Technologies, California, USA) on a BIO-RAD iQ5 iCycler Thermal Cycler (Bio-Rad, California, USA). A total of 20 µL of qPCR reaction mixture was used where each contained 1x master mix, 37.5 nM reference dye, 250 nM forward and reverse primers, RNase/ DNase-free water and 200 ng of the first-strand cDNA. PCR reactions were carried out at 95 °C for 30 min, then by 40 cycles of 5 s at 95 °C and 15 s at 60 °C. Melt curve analysis (55-95 °C with an increment of 0.5 °C every 10 s) were conducted to verify the specificity of each products. A no-template control was included for each pair of primers. The efficiency of PCR amplification was derived from a standard curve generated by a serial dilution with five dilution points of a pool of cDNA samples (refer to section 3.1), ranging from 0.7 to 460 ng (Appendix G). The expression levels of the tested reference genes were determined by Ct values, where Ct is the number of PCR cycles required for the fluorescence to reach the threshold level. Each RT-qPCR reaction was performed in triplicates.

3.13 Reverse Transcription Quantitative Real-Time PCR (RT-qPCR) Analysis

The expressions of two transcripts encoding MC5E (SP01411 and SP02271, respectively), which are involved in the biosynthesis of alginate, two transcripts encoding GDP-D-mannose 4,6 dehydratase (GM46D) (Contig159) and GDP-4-keto-6-

deoxy-D-mannose-3,5-epimerase/4 reductase (GFS) (Contig87), respectively, involved in the biosynthetic pathway of sulfated fucans, in the *S. polycystum* samples, were profiled by using real-time RT-qPCR. The RT-qPCR was conducted in BIO-RAD iQ5 iCycler Thermal Cycler (Bio-Rad) using the Brilliant III Ultra-Fast SYBR® Green QPCR Master Mix (Agilent Technologies). To avoid amplification of unspecific transcripts from multigene family, the forward primers were designed at the open reading frame (ORF) whereas reverse primers were designed at the 3'-untranslated region (UTR). The RT-qPCR reactions were performed in a 96-well thermocycler (Biorad, iQ5) at 95 °C for 3 min, followed by 40 cycles of 5 s at 95 °C and 10 s at 60 °C. Each sample was in triplicate. Non-template control (water template) was also included. The specificity of amplification was checked with a dissociation curve by heating the samples from 55 °C to 95 °C with an increment of 0.5 ºC every 10 s for 81 cycles to confirm the specificity of the PCR amplification, and the absence of primer dimmers (Appendix H).

The relative abundance of transcripts of *S. polycystum* collected on July 2010 was compared to that of the *S. polycystum* collected on December 2010. Quantification of gene expression was conducted by using comparative ΔΔCt method (Hellemans *et al*., 2007) with three housekeeping genes. Ct (cycle threshold) is defined as the number of PCR cycles required for the fluorescence signal to cross the threshold. Three endogenous control transcripts encoding alpha tubulin (Contig6), ribosomal protein L3a (Contig259) and translation elongation factor 1a (Contig62), respectively, from *S. polycystum* were amplified in parallel for normalization. The relative fold change (R) was computed by a formula, $R = E^{-\Delta\Delta Ct}$, where E is the amplification efficiency of the respective genes. The relative expression fold change of dry season sample was calculated by comparing the abundance of transcripts in dry season sample with rainy season sample (control sample). A gene relative transcript abundance if decreased ≥ 50

% or increased ≥ 2-fold compared to that of controls, the gene is regarded as differentially expressed (Rajeevan *et al*., 2001).

Table 3.1: Primers used for RT-qPCR validations

EST or Contig	Abbreviation	Putative annotation	Primer sequence (5'-3')	PCR product length (bp)
SP01411	MC5E-1	Mannuronan C5-epimerase	GGAGCGGCAACATCATTGG CATTCGCCAGACAGTTCATCG	247
SP02271	MC5E-2	Mannuronan C5-epimerase	TCCTCTCCCTTTGCGGCTTGTTGTA TCCTGCCTGACCAGAGTGCCATTC	173
Contig159	GM46D	GDP-D-mannose 4,6-dehydratase	GAGTGCGGAGAACGGATTAAGACA GACAGACGAAGAGGGCGAGAATG	188
Contig87	GFS	GDP-4-keto-6-deoxy-D-mannose epimerase-reductase	GCCCGCTTCGCCAGTTCATTT GCCGTCTCGTAGTTCTCGCAGA	322
Contig6	AT	Alpha tubulin	CGCTTGAGAAGGACTACGAGGAG AGCAACACAGCAGATAACCGAAGA	169
Contig259	RPL3a	Ribosomal protein L3a	CCCCAAGAAACGGACCAAGA GCCCACAACTCCAACAACCA	223
Contig62	TEF1A	Translation elongation factor 1a	CGACAAGAAAGAGAAGTCCAAGAA AGTGTGTTAATGCCGCTTCCCTAC	197

CHAPTER 4: RESULTS

4.1 Optimization of RNA Extraction

In this study, five different RNA extraction methods, that are, Method 1-modified from Hong *et al*. (1997), Method 2-modified from Kim *et al*. (1997), Method 3-modified from Wang *et al*. (2008), Method 4-Wong *et al*. (2007) and Method 5-modified from Wong *et al*. (2007), were used to extract total RNA from *S. polycystum*. These RNA isolation methods were widely used to extract RNA from seaweeds and plants.

Method 1 used PVPP, LiCl and guanidinium thiocyanate in the extraction buffer. This modified RNA extraction method from various seaweed tissues was from Hong *et al*. (1997), and it worked well for *Porphyra* (Hong *et al*., 1995). The low A_{260}/A_{280} ratio (1.175 ± 0.064) and the low A_{260}/A_{230} ratio (0.535 ± 0.064) obtained from method 1 (Table 4.1) in the present study, suggested that guanidinium thiocyanate is not suitable for removing proteins and polysaccharides from *S. polycystum*.

Method 2 used EDTA, EGTA and LiCl in the extraction buffer. It is a modified method from Kim *et al*. (1997). This method improved the RNA yield and quality from *Gracilaria changii* (Chan *et al*., 2004), but it did not work for *S. polycystum* in the present study. Even though the yield of RNA extracted from this method was the highest (142.15 ± 10.536 µg/g) among all the five methods, yet, the extraction method yielded a large amount of undissolved polysaccharides (Figure 4.1, lane 2). Low A_{260nm}/A_{280nm} ratio (1.105 ± 0.035) and A_{260nm}/A_{230nm} ratio (0.655 ± 0.078) obtained from method 2 suggested that this method was not suitable for RNA extraction from *S. polycystum*.

Method 3 is a modified method from Wang *et al*. (2008). The quality of RNA extracted from this method was better than method 1 and method 2 with better A_{260}/A_{280} ratio (1.515 ± 0.064) and A_{260}/A_{230} ratio (1.01 ± 0.085). However, the quality of the extracted RNA was still not satisfactory.

Method 4 was used to extract RNA from *S. binderi*, a brown seaweed (Wong *et al*., 2007). The quality (A_{260}/A_{280} ratio at 1.680 ± 0.057 and A_{260}/A_{230} ratio at 1.955 ± 0.035) and yield (27.07 ± 3.818 µg/g fresh weight) of the RNA were still less satisfactory.

Method 5 is an improved RNA extraction method of Method 4. This protocol eliminated most of the interfering molecules efficiently and yielded translucent and water-soluble RNA pellets. Two distinct 28S and 18S rRNA bands were shown clearly on formaldehyde agarose gel stained with EtBr (Figure 4.2), which indicated that the RNA was intact. The average RNA yield obtained from *S. polycystum* was 43.52 ± 0.716 µg g^{-1} fresh weight. The RNA obtained from method 5 showed no polysaccharide, polyphenolic compounds, and protein contamination (A_{260}/A_{230} absorbance ratio of 2.423 ± 0.115) compared to the other methods (A_{260}/A_{230} absorbance ratio less than 2.0) (Table 4.1).

Table 4.1: Yield and quality of RNA from *S. polycystum* using five methods

Method	Purity assessment[a]		Total RNA (µg)	Yield (µg/g FW)
	A_{260}/A_{280}	A_{260}/A_{230}		
1. Modified from Hong *et al*. (1997)	1.175 ± 0.064	0.535 ± 0.064	131.33 ± 5.197	81.55 ± 3.465
2. Modified from Kim *et al*. (1997)	1.105 ± 0.035	0.655 ± 0.078	213.23 ± 15.803	142.15 ± 10.536
3. Modified from Wang *et al*. (2007)	1.515 ± 0.064	1.01 ± 0.085	61.40 ± 1.704	40.93 ± 1.133
4. Wong *et al*. (2007)	1.680 ± 0.057	1.955 ± 0.035	40.61 ± 5.728	27.07 ± 3.818
5. Modified from Wong *et al*. (2007)	2.039 ± 0.019	2.423 ± 0.115	65.28 ± 1.075	43.52 ± 0.716

FW: fresh weight

[a.] Mean of three samples (standard deviation)

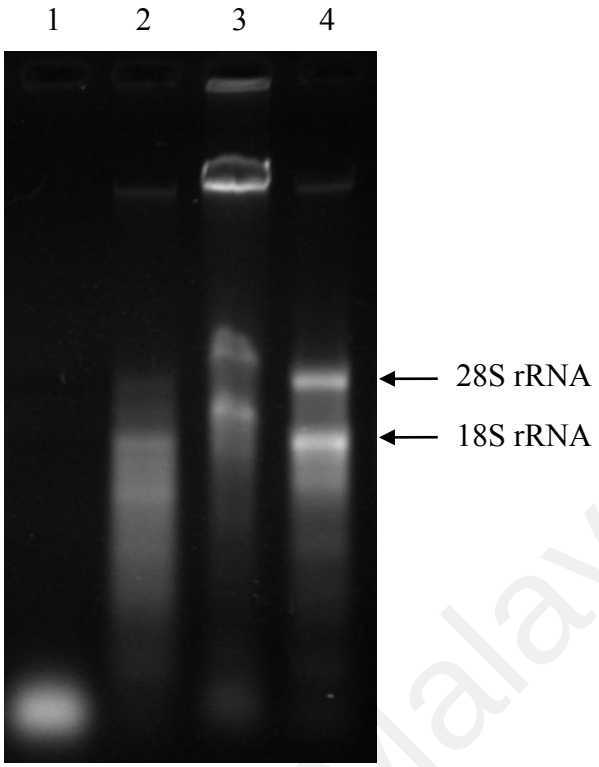


Figure 4.1: RNA of *S. polycystum* isolated using 4 different methods.

Lane 1: Method 1 - modified from Hong *et al*. (1997), lane 2: Method 2 - modified from Kim *et al*. (1997), lane 3: Method 3 – modified from Wang *et al*. (2008), lane 4: Method 4 – Wong *et al*. (2007). The 28S and 18S ribosomal RNA are indicated by arrows. These four RNA extraction methods failed to isolate intact RNA or produced negligible quantity of RNA

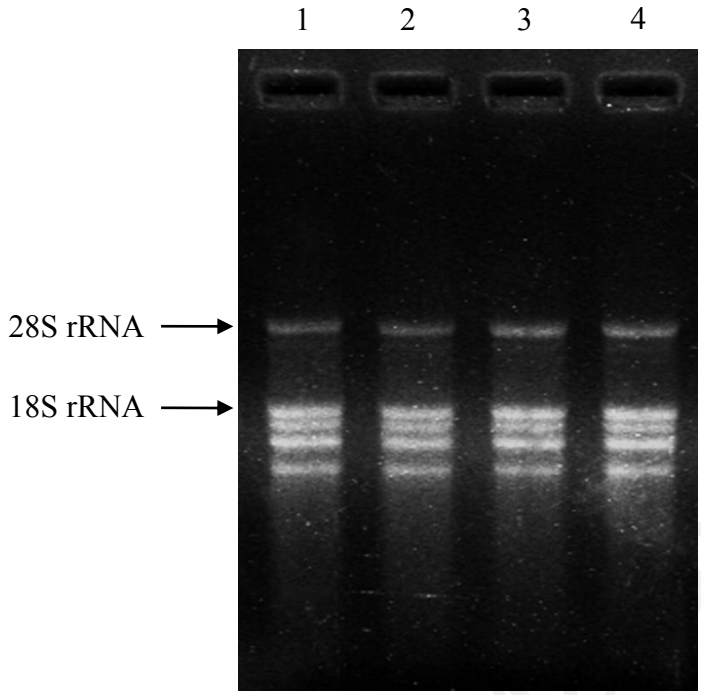


Figure 4.2: Formaldehyde-agarose gel electrophoresis of total RNA (2 µg) from *S. polycystum*, by using Method 5.

4.2 mRNA Isolation

A total of 7.5 µg of mRNA was isolated from ~750 µg of total RNA. The recovery rate of mRNA was between 1 – 2 % (Figure 4.3) in this study in accordance to the estimated recovery rate of mRNA of the kit (1 - 5 %) when fresh intact RNA was used for mRNA isolation.

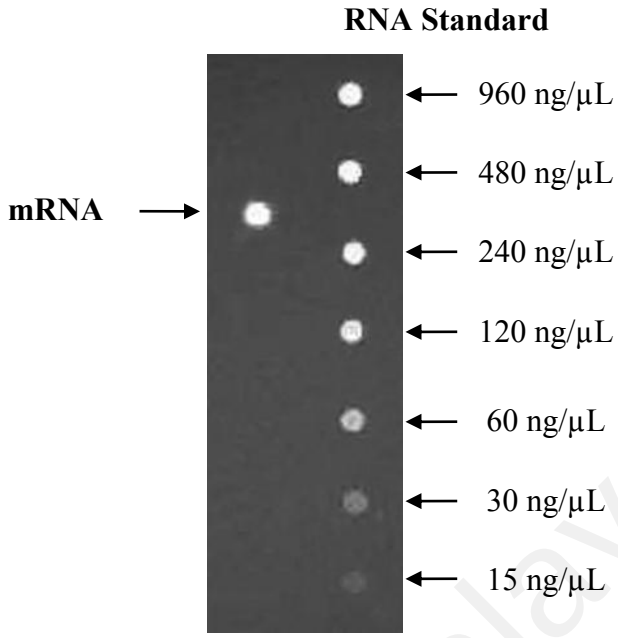


Figure 4.3: Estimation of mRNA concentration on EtBr plate.

A serial dilution (960, 480, 240, 120, 60, 30 and 15 ng/µL) of RNA standard was spotted on an EtBr-stained 1 % (w/v) agarose gel.

4.3 cDNA Library Construction

To construct a good cDNA library, sufficient amount of good quality mRNA was required. A total of 5.6 µg of mRNA was used for construction of cDNA library of *S. polycystum*. After the second strand cDNA synthesis, gel electrophoresis was carried out to estimate the size range of the first and second strand cDNA (Figure 4.4). Both first and second strand cDNA were ranged from 247 bp and above.

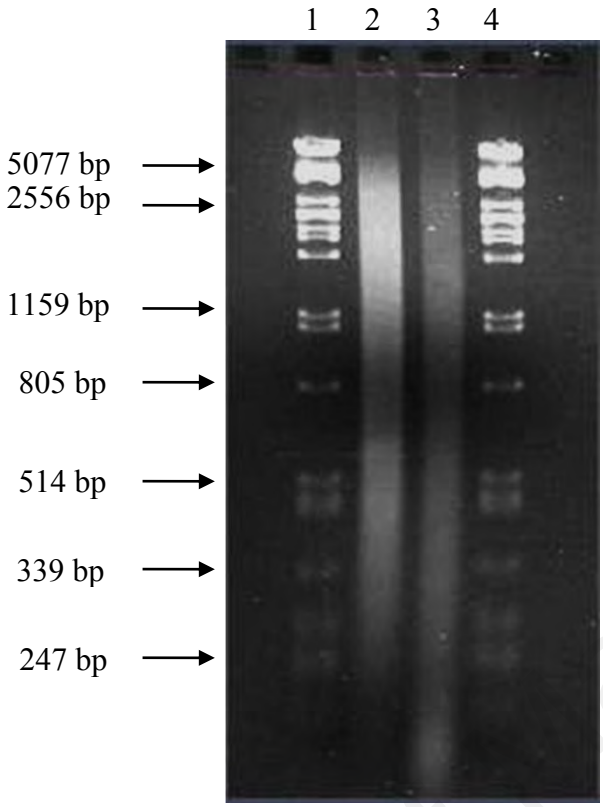


Figure 4.4: First-strand cDNA of *S. polycystum*.

Four percent of the first-strand cDNA synthesis reaction (lane 2) and ~5.5 % of the second-strand cDNA synthesis reaction (lane 3) were electrophoresed on 1 % agarose gel, λ/PstDNA ladder (0.6 µg) was used as a marker (lane 1 and lane 4).

Size fractionation of cDNA was performed to reduce the occurence of smaller inserts (< 500 bp) in the cDNA library (Figure 4.5). Small inserts (< 500 bp) may produce sequencing results that are not informative for functional annotation. A good primary cDNA library consisting of 1.0 x 10^6 clones. It was revealed by X-Gal/IPTG screening which constructed with around 99 % of recombinant clones. The amplified library was 4.2 x 10^{10} pfu/mL. A total of 23 colonies were selected randomly from the primary cDNA library for PCR amplification and restriction enzyme (RE) verification to check the percentage of colonies with insert. Figure 4.6 showed that all the colonies were with insert and the insert size ranged from 500 bp to 4000 bp.

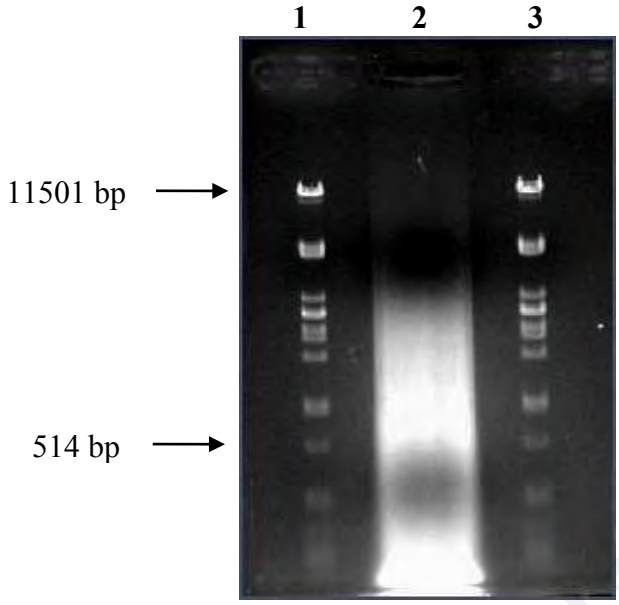


Figure 4.5: Double stranded cDNA before size fractionation.

After digested with *Xho* I, the cDNA library was electrophoresed on 1 % agarose gel. cDNA within the range of 500 bp to 10,000 bp was excised (lane 2). The λ/Pst DNA ladder (0.6 µg) was used as marker (lane 1 and lane 3).

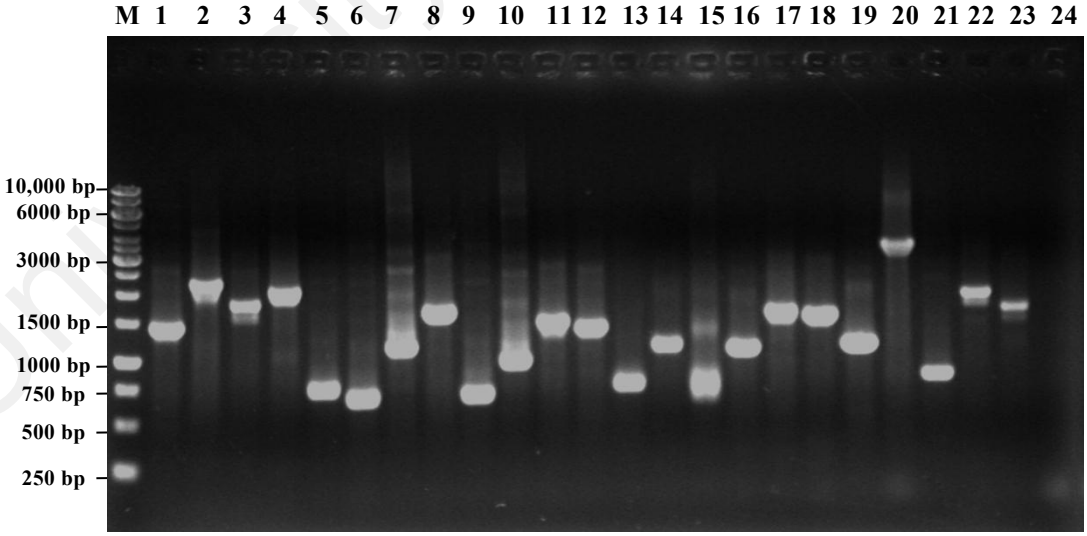


Figure 4.6: PCR product of cDNA clones.

PCR amplification was performed to verify cDNA inserts from 23 randomly chosen clones of *S. polycystum* cDNA library using T3 and T7 primers (lanes 1–23) and negative control (lane 24). Lane M, 1 kb DNA Ladder (Fermentas, Ontario, Canada), lanes 1–23, insert cDNA clones of 600 bp–4 kb in length, lane 24, negative control.

4.4 Generation of ESTs

Single-pass 5' sequencing of 2586 cDNA clones resulted in 2577 high-quality readable EST sequences with an average read length of about 878bp. To reduce redundancy, these ESTs were assembled into 295 contigs and 1429 singletons, by using CAP3, giving a total of 1724 unigenes (Table 4.2). About 244 contigs (83 %) out of 295 contigs consisted of two to five ESTs, 48 contigs (16 %) consisted of six to 19 ESTs, whereas only three contigs (1 %) consisted of more than 20 ESTs. The singletons and contigs ranged from 117 bp to 1131 bp (average length of 829 bp) and 475 bp to 3005 bp (average length of 1119 bp), respectively. The redundancy of *S. polycystum* ESTs was estimated to be 45 %.

Table 4.2: Summary of *Sargassum polycystum* ESTs based on their matches to GenBank database

Features of ESTs	Number	Percentage
Total ESTs sequenced	2586	-
Number of EST sequences with readable sequence	2577	99[a]
Redundant sequences	1148	45[b]
Number of unigenes	1724	65[b]
Number of singletons	1429	83[c]
Number of contigs	295	17[c]
Average contig size (bp)	1119	-
Longest contig (bp)	3005	-
Longest singleton (bp)	1131	-
Unigenes matched to NCBI non-redundant database	1142	66.2[c]
Unigenes matched to Swiss-Prot database	798	46.3[c]

[a]Percentage of total EST sequenced. [b]Percentage of total readable ESTs. [c]Percentage of number of unigenes

4.5 Functional Annotation

Of 1724 sequences, 798 (46.3 %) sequences and 1142 (66.2 %) sequences have significant matches with sequences in Swiss-Prot and NCBI non-redundant (nr) protein

database, respectively (Table 4.2). The *S. polycystum* unigenes were ranked according to the number of contributing ESTs in Table 4.3. The vast majority of the highly represented transcripts are ribosomal proteins based on Swiss-Prot protein database. Eleven of these unigenes matched to ribosomal proteins. Based on NCBI nr protein database, majority of the highly represented transcripts are conserved unknown protein that share homology to a brown seaweed, *Ectocarpus siliculosus*. Other unigenes had significant matches to actin (contig164), heat shock protein (contig48), short-chain dehydrogenase (contig187), annexin a3 (contig210), aquaporin (contig252) and elongation factor protein (contig15), respectively. Contig164 had significant matches to actin that shares homology to brown seaweeds (*Fucus vesiculosus* and *Saccharina japonica*), and contig48 had significant matches to heat shock protein from *Fucus serratus* (brown seaweed) and *Arabidopsis thaliana* (higher plant), respectively. However, three unigenes (contigs9, contig130 and contig206) of *S. polycystum* which are highly abundant in this library had no significant similarity to any sequences in the Swiss-Prot and NCBI non-redundant (nr) protein database, respectively (Table 4.3). Since NCBI nr protein database provides more information, it was used for subsequent annotations including gene ontology mapping and KEGG analysis. The top hits of the ESTs to NCBI nr protein database showed matches to sequences from algal sequences (92 %) (Figure 4.7). A total of 1623 EST sequences (86.1 %) matched with brown algal sequences from *E. siliculosus*, *F. serratus*, *F. vesiculosus*, *F. distichus*, *L. digitata*, *Saccharina japonica*, *Saccharina latissima*, *Ascophyllum nodosum*, *S. binderi*, *S. fusiforme*, *S. horneri*, *S. integerrimum*, *S. henslowianum* and *Scytosiphon lomentaria*. A total of 79.8 % ESTs showed high similarity to *E. siliculosus*, followed by *Saccharina japonica* (2 %). Other algae (5.9 %) including Rhodophyta, Chlorophyta, Crytophyta, Haptophyta, cyanobacteria, miozoa and diatom. Some of the ESTs showed high similarity to animal sequences (2.6 %), i.e. fish, bird, insect, primate. A low number of

the ESTs showed high similarity to plant (2.1 %). Other ESTs were similar to fungal sequences (1 %), bacterial sequences (1.1 %), protozoa sequences (1.1 %) and viral sequences (0.1 %).

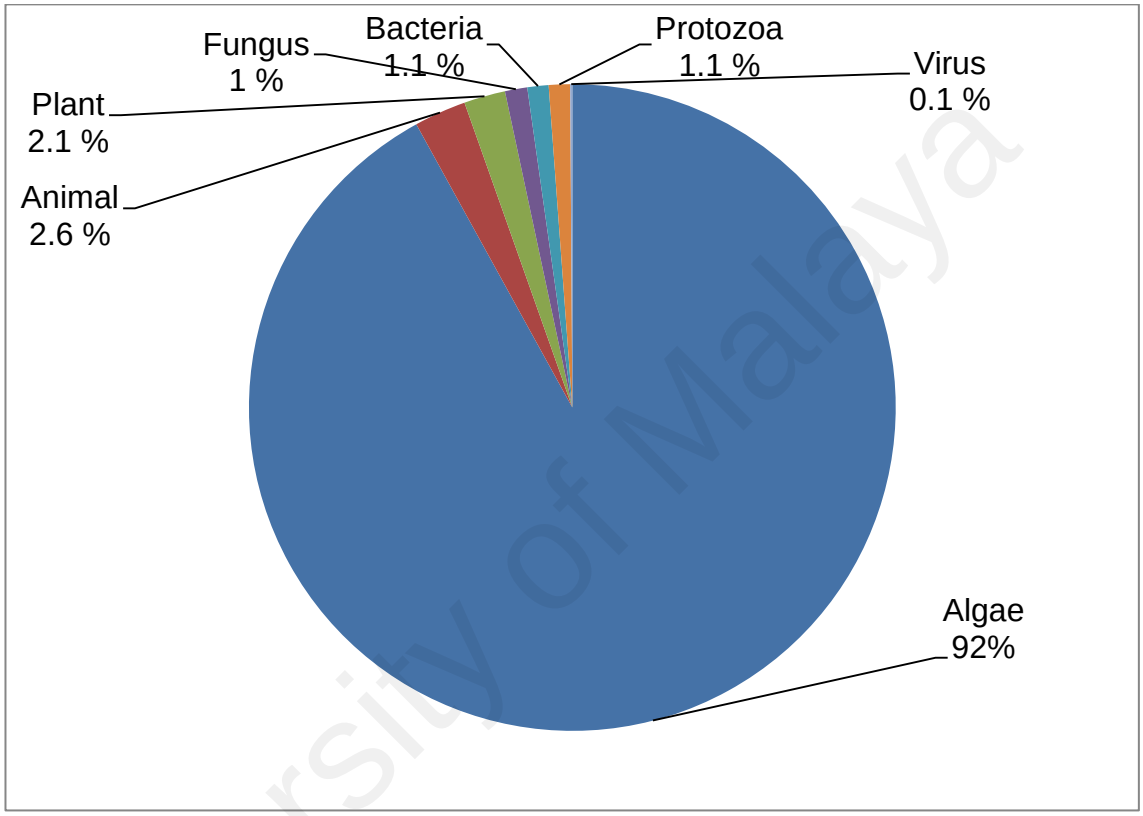


Figure 4.7: ESTs with homology to sequences from various organisms.

Most of the ESTs showed significant homology to algae species (92 %) based on NCBI non-redundant (nr) database

.

Table 4.3: The top 25 unigenes of *S. polycystum*

Accession no.	No. of ESTs	Sequence length	Sequence description match to NCBI database	E-value	% identity	Sequence description match to Swiss-Prot	E-value	% identity
Contig62	30	2143	eukaryotic translation elongation factor 1 alpha 1 [*Ectocarpus siliculosus*]	0	89.80	elongation factor 1-alpha 1 [rat]	0	85.95
Contig164	30	1608	actin [*Saccharina japonica*]	0	96.55	actin [*Fucus vesiculosus*]	0	93.60
Contig9	22	828	---NA---			---NA---		
Contig48	17	1056	heat shock protein hsp20 [*Fucus serratus*]	3.31E-79	75.98	22 kDa heat shock protein [*Arabidopsis thaliana*]	3E-17	58.75
Contig224	17	1678	conserved unknown protein [*Ectocarpus siliculosus*]	5.28E-139	82.08	60s acidic ribosomal protein p0 [soybean]	4E-92	52.57
Contig3	16	1477	conserved unknown protein [*Ectocarpus siliculosus*]	5.26E-105	83.16	40s ribosomal protein s15a [human]	1.64E-69	85.75
Contig32	14	1568	40s ribosomal protein s27 [*Ectocarpus siliculosus*]	8.90E-48	90.95	40s ribosomal protein s27-2 [*Arabidopsis thaliana*]	2.67E-38	82.50
Contig134	13	1910	conserved unknown protein [*Ectocarpus siliculosus*]	1.90E-150	88.85	40s ribosomal protein s3-1 [*Arabidopsis thaliana*]	2.75E-114	78.30
Contig71	12	2197	hypothetical protein Esi_0036_0141 [*Ectocarpus siliculosus*]	1.73E-07	63.00	- NA-		
Contig194	12	3005	cell wall-associated hydrolase [legume]	2.42E-96	77.65	Uncharacterized protein ORF91 [orchid]	1.47E-40	82.50
Contig184	11	1859	conserved unknown protein [*Ectocarpus siliculosus*]	2.92E-122	84.90	40s ribosomal protein s7 [rat]	5.95E-73	75.75
Contig187	11	1555	putative short-chain dehydrogenase [*Ectocarpus siliculosus*]	2.13E-120	67.65	l-xylulose reductase [cattle]	8.12E-38	40.20

NA = not available

Table 4.3 (continued)

Accession no.	No. of ESTs	Sequence length	Sequence description match to NCBI database	E-value	% identity	Sequence description match to Swiss-Prot	E-value	% identity
Contig206	11	935	---NA---			---NA---		
Contig82	10	906	conserved unknown protein [*Ectocarpus siliculosus*]	1.97E-61	81.85	40s ribosomal protein s10 [moth]	2.20E-38	76.95
Contig84	10	1160	cell wall-associated hydrolase [legume]	5.89E-65	66.00	---NA---		
Contig126	10	1684	conserved unknown protein [*Ectocarpus siliculosus*]	3.49E-143	83.05	60s ribosomal protein l10 [fly]	6.84E-99	80.20
Contig252	10	1975	aquaporin [*Ectocarpus siliculosus*]	7.72E-85	56.95	aquaporin pip2-1 [*Arabidopsis thaliana*]	3.55E-44	53.95
Contig26	9	952	40s ribosomal protein s4 [*Ectocarpus siliculosus*]	0	94.27	40s ribosomal protein s4 [chicken]	3.58E-83	62.30
Contig59	8	992	conserved unknown protein [*Ectocarpus siliculosus*]	2.63E-142	76.65	60s ribosomal protein 17a [*Arabidopsis thaliana*]	1.46E-83	62.30
Contig75	8	1264	conserved unknown protein [*Ectocarpus siliculosus*]	9.03E-158	85.65	40s ribosomal protein s0-a [yeast]	5.36E-102	82.35
Contig130	8	901	-NA-			- NA -		
Contig153	8	1009	conserved unknown protein [*Ectocarpus siliculosus*]	4.24E-146	82.45	40s ribosomal protein s3a [legume]	1.18E-105	77.60
Contig158	8	1157	conserved unknown protein [*Ectocarpus siliculosus*]	2.47E-12	70.50	- NA -		
Contig210	8	1839	annexin x [*Ectocarpus siliculosus*]	2.42E-83	63.67	annexin a3 [rat]	1.28E-23	52.40
Contig15	7	1815	elongation factor EF-3 [*Ectocarpus siliculosus*]	0	62.30	elongation factor 3b [yeast]	7.37E-79	56.40

Of the 1142 unigenes with significant matches to NCBI nr protein database, 695 (~40.3 %) were classified into one or more GO categories, resulting in a total of 2998 GO terms (Appendix E). A total of 1029 (59.7 %) unigenes could not be mapped to any GO terms. These GO terms were divided into three principal categories, Biological Process (BP) (43 %), Molecular Function (MF) (36 %) and Cellular Component (CC) (21%), with the majority in the BP and MF categories. The distribution of the gene functions in each of these three categories is shown in Figure 4.8A-C, respectively. Among the genes in the BP category (Figure 4.8A), most GO terms at level 3 were related to the ‗organic substance metabolic process' (GO:0071704, 16 %), followed by ‗primary metabolic process' (GO:0044238, 15%) and ‗cellular metabolic process' (GO:0044237, 15 %). In the MF category, the majority of the GO terms were associated with ‗organic cyclic compound binding' (GO:0097159, 17 %) and ‗heterocyclic compound binding' (GO:1901363, 17 %) (Figure 4.8B**)**. The categories ‗intracellular' (GO:0005622, 20 %) and ‗intracellular part' (GO:0044424, 19 %) were the two most abundant GO terms in CC category (Figure 4.8C). The term ‗regulation of cellular process' (GO:0050794, 3 %) was the least abundant GO in the BP category, whereas in the MF category, ‗transferase activity' (GO:0016740, 5 %) was the least abundant GO term. While in the CC category, ‗membrane protein complex' (GO:0098796, 3 %) was the least abundant GO term.

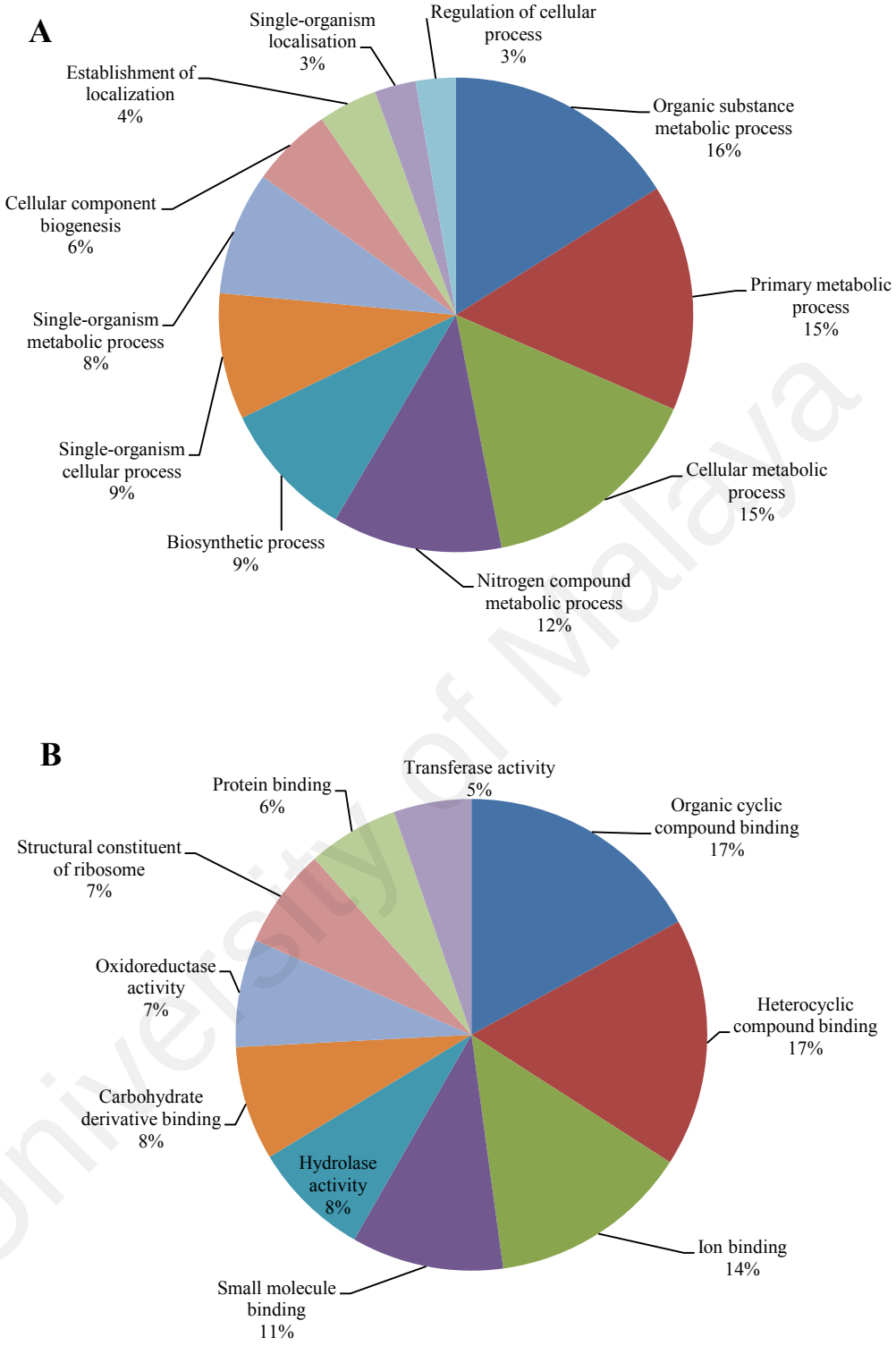


Figure 4.8: Functional classifications of ESTs from *S. polycystum*.

The GO terms were deduced from NCBI nr protein database BLAST hits of the *S. polycystum* cDNA library. Three GO categories are presented: (A) biological process (ontology level 3), (B) molecular function (ontology level 3), and (C) cellular component (ontology level 3). The percentage of sequences is shown for each functional category.

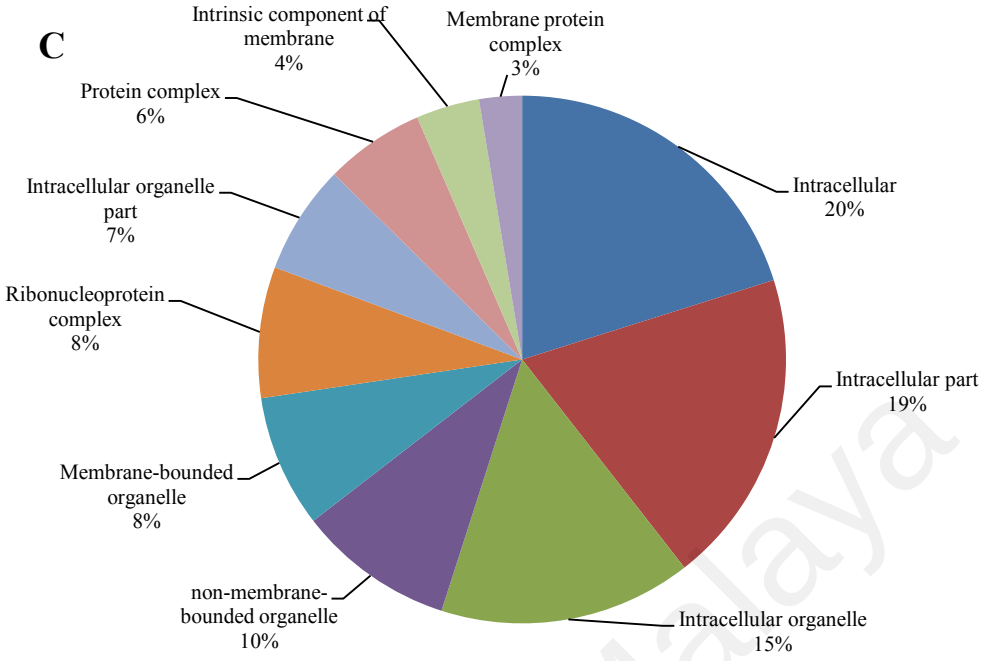


Figure 4.8, continued.

Of the 1724 unigenes, InterPro was able to map 853 unigenes of *S. polycystum* to different databases (Table 4.4). InterPro intergrates the signatures, capitalizing on the respective strengths of the individual databases by classifying sequences into protein families and to predict the presence of important domains and sites (Mitchell *et al*., 2014).

Table 4.4: InterPro intergrates signatures of different databases for classifying sequences of *S. polycystum* into protein family, to predict important domains and sites.

Database	Number of unigenes
Signal Peptide (SP)	199
Transmembrane helices (TMD/TMHMM)	375
GENE3D	505
IPR	630
PFAM	573
PANTHER	571
SUPERFAMILY	521
PRINTS	110
PROSITE_PROFILE	193
SMART	135
PRODOM	18

The 25 most frequent InterPro protein domains and important sites among *S. polycystum* unigenes are presented in Table 4.5. The most abundant protein domains found in the *S. polycystum* unigenes were the P-loop containing nucleoside triphosphate hydrolase superfamily, NAD(P)-binding Rossman-fold domains, thioredoxin-like fold domain, protein kinase-like domain, protein kinase, catalytic domain and glucose/ ribitol dehydrogenase family. Among these, three protein domains (IPR011009, IPR000719, G3DSA:3.30.200.20) are related to protein kinase, while three protein domains (IPR012336, IPR010987, IPR017933) are related to glutathione S-transferase, and two protein domains/family (IPR002347, IPR002198) are related to short-chain dehydrogenase.

Table 4.5: Most abundant InterPro functional domains and important sites found among the unigenes of *S. polycystum*

Interpro ID	Description	No. of unigenes
SSF52540(SUPERFAMILY)	P-loop containing nucleoside triphosphate hydrolase superfamily	37
G3DSA:3.40.50.300	P-loop containing nucleotide triphosphate hydrolases	36
SSF51735(SUPERFAMILY)	NAD(P)-binding Rossman-fold domains superfamily	26
IPR016040	NAD(P)-binding domain (D)	25
IPR012336	Thioredoxin-like fold (D)	20
IPR011009[1]	Protein kinase-like domain(D)	11
IPR000719	Protein kinase, catalytic domain (D)	10
IPR002347[2]	Glucose/ ribitol dehydrogenase (F)	10
SSF54236(SUPERFAMILY)	Ubiquitin-like superfamily	10
G3DSA:3.10.20.90	Phosphatidylinositol 3-kinase catalytic subunit: chain A, domain 1	9
IPR002198	Short-chain dehydrogenase/reductase SDR (Family)	9
IPR010987	Glutathione S-transferase, C-terminal-like (D)	9
G3DSA:2.40.30.10	Translation factors	8
G3DSA:3.30.200.20	Phosphorylase kinase; domain 1	8
G3DSA:3.40.50.150	Vaccinia virus protein VP39	8
IPR000626	Ubiquitin (D)	8
IPR009000	Traanslation elongation/ initiation factor/ Ribosomal, beta-barrel	8
IPR013785	Aldolase-type TIM barrel	8
IPR015943	WD40/ YVTN repeat-like-containing domain (D)	8
IPR017933	Glutathione S-transferase/chloride channel, C-terminal (D)	8
IPR019955	Ubiquitin supergroup	8
PTHR23115 (PANTHER)	Elongation factor 1-alpha 1	8
SSF51905 (SUPERFAMILY)	FAD/NAD(P)-binding domain	8
SSF53335(SUPERFAMILY)	S-adenosyl-L-methionine-dependent methyltransferases	8
SSF53474(SUPERFAMILY)	Alpha/ beta-hydrolases	8

[1]Contained in IPR000719

[2]Contained in IPR002198

Of the 1724 unigenes, only 205 unigenes were mapped to 127 Enzyme Commission (EC) code including oxidoreductases (n = 66), transferases (n = 43), hydrolases (n = 44), lyases (n = 18), isomerases (n = 19) and ligases (n = 19), where n is the number of unigenes in the EC categories. ECs with the highest number of unigenes include EC:3.4.25.0 for threonine endopeptidases (n = 9) and EC:5.2.1.8 for peptidylprolyl isomerase (n = 9).

KEGG pathways of the unigenes of *S. polycystum* were analysed. A total of 277 unigenes were assigned to 64 biochemical pathways, of which 61 are metabolism pathways. Genetic information processing (translation); environmental information processing (signal transduction); and organismal system (immune system) were represented by one pathway each. For metabolism pathway, carbohydrate metabolism is the majority sub-group, followed by global and overview map, energy metabolism, etc. (Appendix F). The 11 most abundant KEGG pathways of the *S. polycystum* unigenes are shown in Table 4.6.

Table 4.6: KEGG pathways with most abundant unigenes of the *S. polycystum* cDNA library derived from NCBI nr annotations.

KEGG pathway	No. of unigenes	No. of enzyme type
Purine metabolism	75	8
Thiamine metabolism	68	1
Biosynthesis of antibiotics	52	43
Oxidative phosphorylation	22	6
Carbon fixation in photosynthetic organisms	17	11
Glycolysis/ Gluconeogenesis	16	10
Fructose and mannose metabolism	16	8
Aminobenzoate degradation	12	2
Glutathione metabolism	10	6
Pentose phosphate pathway	10	6
Citrate cycle (TCA cycle)	9	10
T cell receptor signaling pathway	9	1

Table 4.7 showed the putative transcripts encoding for enzymes of selected metabolic pathways, among these are enzymes related to stress responses and enzymes for fructose and mannose metabolism.

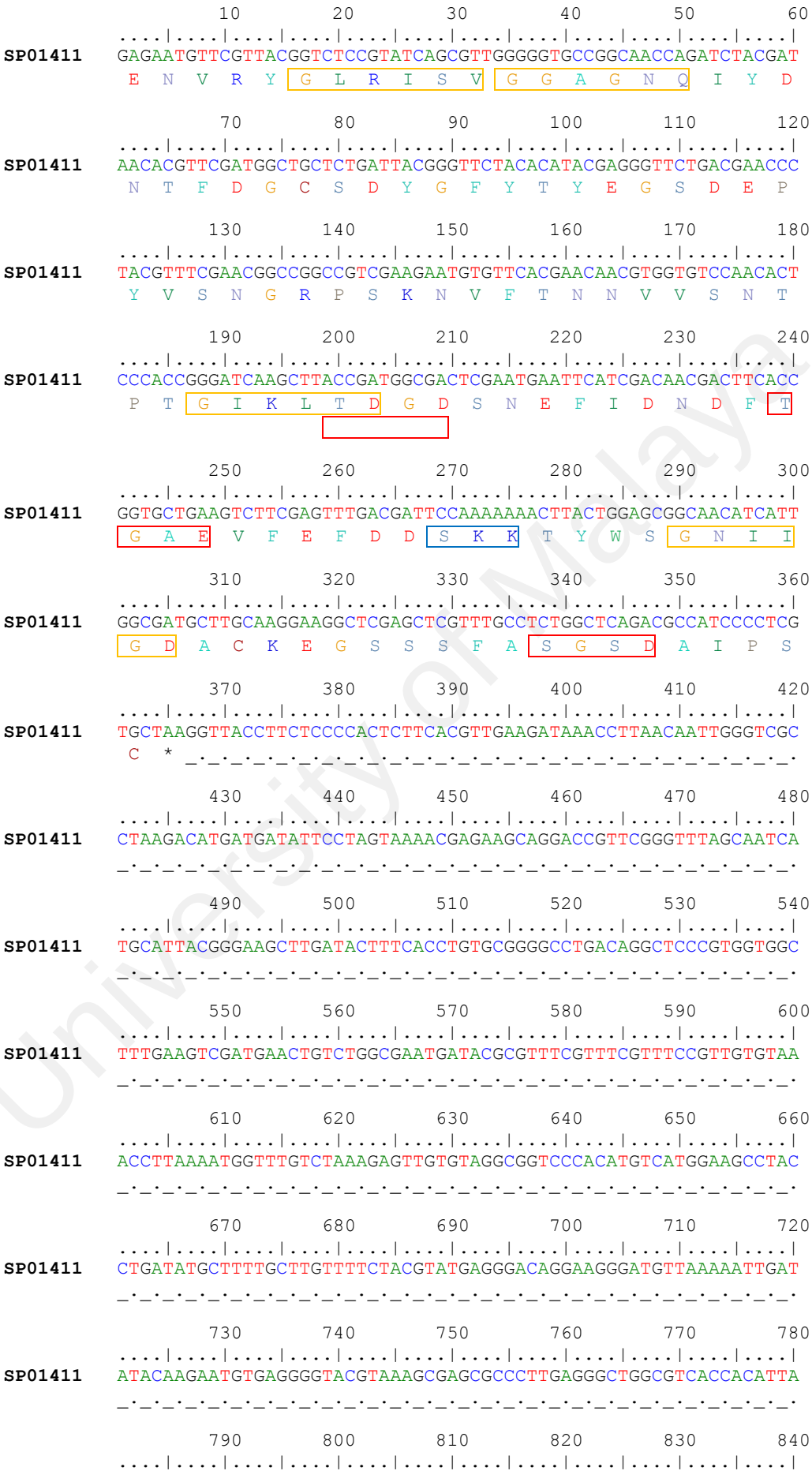
Table 4.7: Putative transcripts of *S. polycystum* encoding for enzymes of selected pathways

Pathways structural enzymes	No. of ESTs	EST or Contig	*E*-value
Fructose and mannose metabolism			
GDP-L-fucose synthase; gdp-4-keto-6-deoxy-D-mannose-3,5-epimerase-4-reductase	3	Contig87	1.25E-144
Fructose-bisphosphate aldolase	5	Contig170	2.88E-144
Fructose-bisphosphate aldolase	1	SP01103	5.91E-100
Phosphomannomutase	4	Contig57	5.99E-121
GDP-mannose dehydratase	3	Contig159	4.67E-165
GDP-mannose dehydratase	3	Contig254	0
Ribose-5-phosphate isomerase	1	SP02422	1.97E-65
Triosephosphate isomerase	1	SP01238	1.30E-45
Triosephosphate isomerase	1	SP01281	1.20E-146
Glyceraldehyde-3-phosphate dehydrogenase	1	SP01272	5.70E-66
UDP-glucose 6-dehydrogenase	2	Contig211	0
GDP-mannose 6-dehydrogenase	1	SP01705	3.70E-52
GDP-mannose 6-dehydrogenase	1	SP00900	1.40E-54
Mannitol 1-phosphate dehydrogenase	1	SP01050	4.50E-131
Mannitol 1-phosphate dehydrogenase	4	Contig109	9.60E-119
Mannitol 1-phosphate dehydrogenase	1	SP01780	2.00E-81
Stress related/ Defense/ Detoxification			
Ascorbate peroxidase	2	Contig11	2.37E-20
Cytochrome c peroxidase	1	SP02475	8.84E-23
Glutamine synthetase	1	SP00780	3.13E-164
Glutathione S-transferase	5	Contig103	1.27E-26
Glutathione reductase	2	contig196	2.20E-103
Glutathione peroxidase	1	SP02517	2.19E-36
Superoxide dismutase	2	Contig191	1.85E-32
Heat shock protein 20 (HSP 20)	33	Contig48, Contig 77, Contig205,Contig274, SP00423, SP00680, SP02397, SP02572	9.39E-79
Heat shock protein 40 (HSP 40)	3	SP00533, SP02368, SP02369	2.53E-40
Heat shock protein 60 (HSP 60)	2	Contig279	9.18E-87
Heat shock protein 70 (HSP 70)	6	SP00097, SP00817, SP00910, SP01721, SP02363, SP02550	9.57E-90
Heat shock protein 90 (HSP 90)	3	SP00548, SP01543, SP01695	2.44E-94

4.6 Sequence Analyses

4.6.1 Sequence Analysis of Mannuronan C5-Epimerase (MC5E)

In this study, three partial transcripts (SP01411, SP01750 and SP02271) encoding putative mannuronan C5-epimerase were identified from the EST of *S. polycytsum*. However, frameshift mutation has occurred in transcript SP01750, resulting in a completely different translation from the original. Thus, transcript SP01750 was not used for further analysis. The transcript SP01411 is 1793 bp length, after sequencing using universal primers T3 and T7. The EST sequence of SP01411 contained a partial coding region, stop codon with poly(A)$^+$ tail and long 3' untranslated region (3' UTR), but it did not have the start codon and most of the coding sequence (cds) when compared to the transcripts encoding mannuronan C5-epimerase from *L. digitata*, *E. siliculosus*, *S. japonica* (Figure 4.11). Nonetheless, the amino acid sequence encoding by SP01411 matches to three known protein profiles from the PROSITE database, including N-myristoylation site (G - {EDRKHPFYW}- x(2) – [STAGCN] – {P} where G is the N – myristoylation site), protein kinase C phosphorylation site ([ST] – x - [RK] where S or T is the phosphorylation site), and casein kinase II phosphorylation site ([ST] – x(2) – [DE] where S or T is the phosphorylation site) (Figure 4.9).



```
SP01411   GACATGATGTTGAGGATATCGCTGCCATGTTTGCCCTGTTCGAGCAGGGGGCTCGCGTGC
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·

                 850       860       870       880       890       900
          ....|....|....|....|....|....|....|....|....|....|....|....|
SP01411   GGACGGTCCTTTGCATCTGTGAGTTGTATGGCCTATCTAGGAAGAGACAAATATACCTTA
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·

                 910       920       930       940       950       960
          ....|....|....|....|....|....|....|....|....|....|....|....|
SP01411   GATGTGAGTATTGCATTGCTGCTGCTAGCGTTGTTTTTGGGACGCTAAGCGAATGCCCTT
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·

                 970       980       990      1000      1010      1020
          ....|....|....|....|....|....|....|....|....|....|....|....|
SP01411   TATCGGTCCGTCTTGTCTTGTTTGCTTTTTCTTCATACTGTTTGGTTGCTCTTTAGCAGA
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·

                1030      1040      1050      1060      1070      1080
          ....|....|....|....|....|....|....|....|....|....|....|....|
SP01411   GCGAGTACTTTTCTCTGTGGTTCCTTGGACCCATCGGGATCCAGATAAAATAAAGTGGGC
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·

                1090      1100      1110      1120      1130      1140
          ....|....|....|....|....|....|....|....|....|....|....|....|
SP01411   ATTTTATATTGAATTGAAGCAGTAGACGTGCGCTTTACCGATTTCAAGAGAGTCTTTTGT
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·

                1150      1160      1170      1180      1190      1200
          ....|....|....|....|....|....|....|....|....|....|....|....|
SP01411   CGGTGGTGCGATCATGTTTCTGTGAAGAGTTTTGATCGTCAAGGATTTCACTTATGTTAT
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·

                1210      1220      1230      1240      1250      1260
          ....|....|....|....|....|....|....|....|....|....|....|....|
SP01411   GGAGCAAGTTTAGCCCCGGTACTACTTGGCTTTTAAAGCCTGCCGCCTGCGAGTAGTGAA
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·

                1270      1280      1290      1300      1310      1320
          ....|....|....|....|....|....|....|....|....|....|....|....|
SP01411   GTATTTTTCAGAAGCAATATCAAGTGTGTGGTCAGATGTTTCGCCCTCATCTCGATCGCG
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·

                1330      1340      1350      1360      1370      1380
          ....|....|....|....|....|....|....|....|....|....|....|....|
SP01411   CTACTGCAGAGCGTTGTAAAGTAAATATATATTAGTTGTTTTTCACTCAAGATGTTGGAC
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·

                1390      1400      1410      1420      1430      1440
          ....|....|....|....|....|....|....|....|....|....|....|....|
SP01411   ACGAGACTACTCGATAATGTCACGTTGTAGCTCAGACAGCCTCTGCTTTCACTGACTTGC
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·

                1450      1460      1470      1480      1490      1500
          ....|....|....|....|....|....|....|....|....|....|....|....|
SP01411   TGGCTTTCGAAACGTTCGGTAGTTTATACCGATTTCTATGTGAGTGTAGGTAGTATGTGT
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·

                1510      1520      1530      1540      1550      1560
          ....|....|....|....|....|....|....|....|....|....|....|....|
SP01411   TCTGTAAGCCTCGTTTGGAGGCACATTATACAAAGTATATAGCATACCCGCTGTGCTATT
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·

                1570      1580      1590      1600      1610      1620
          ....|....|....|....|....|....|....|....|....|....|....|....|
SP01411   AAGGAAACAAGATCACGGTTTTGTTCTAGGGTTAGCAGCAGTAAGTCTCTAGTAGCACAA
          _·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·_·
```

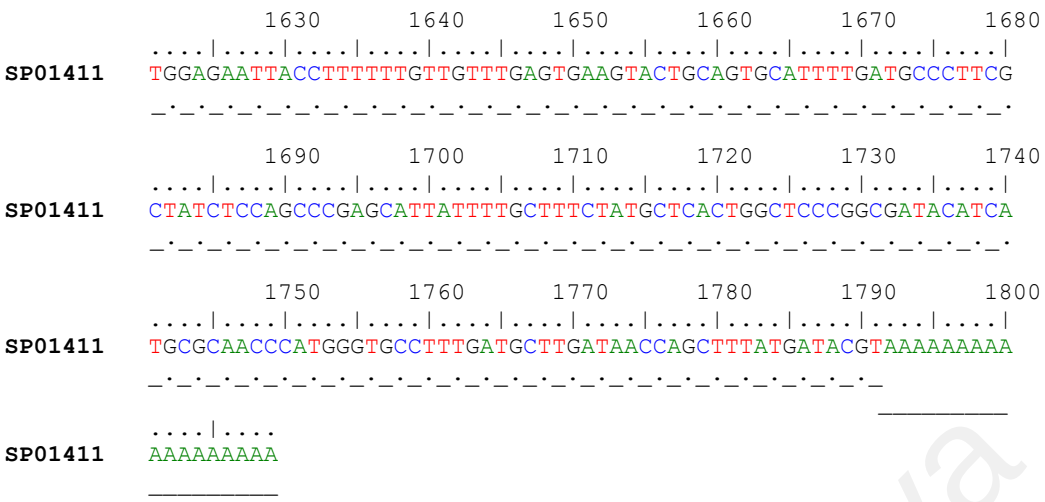
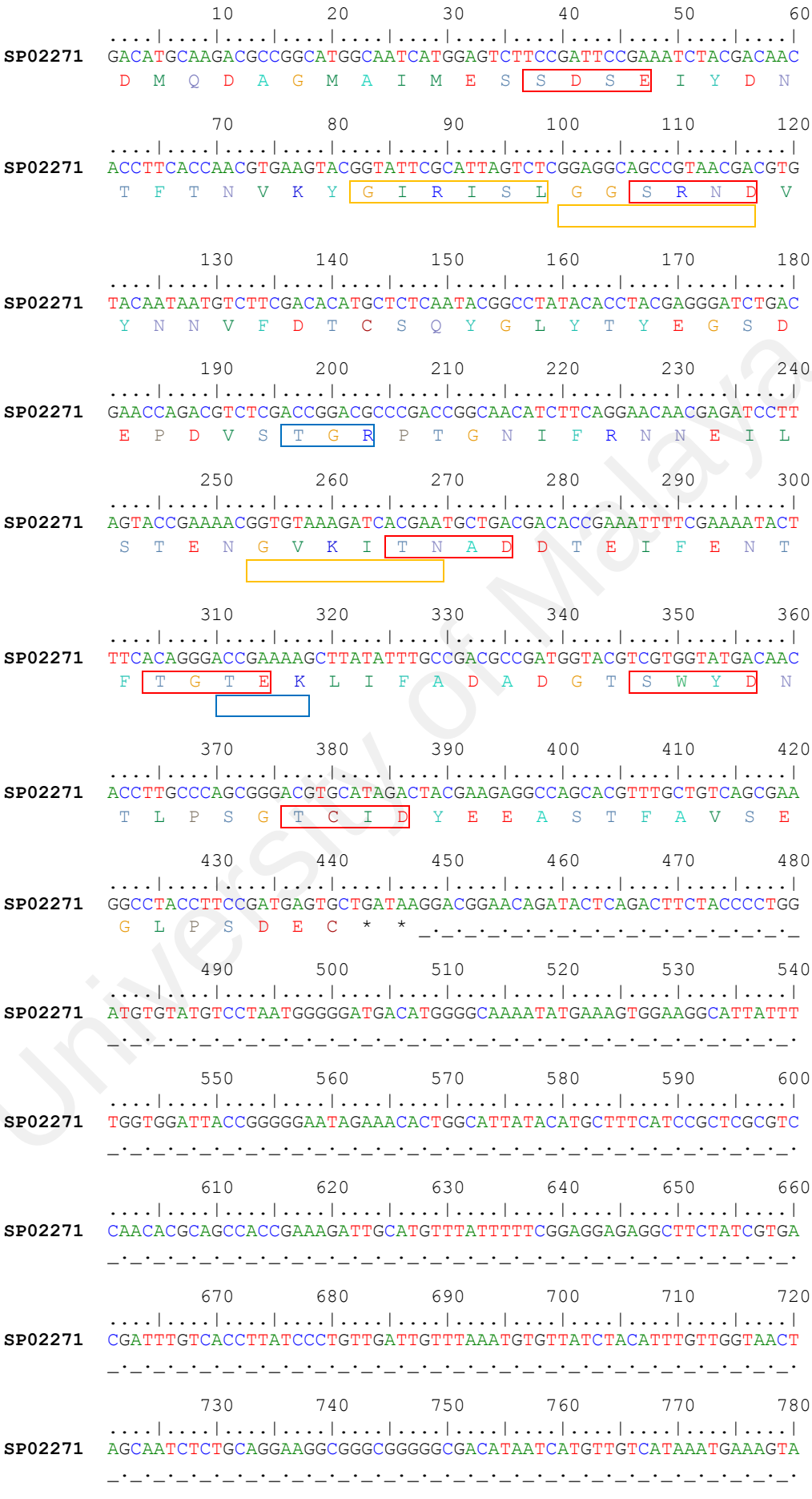


Figure 4.9: Partial transcript sequence encoding mannuronan C5-epimerase transcript (MC5E-1) (SP01411).

The predicted amino acid sequence of the transcript is shown under the nucleotide. _._. indicates the 3' UTR and ___ indicates poly(A)+ tail. For Prosite Scan, ☐ indicates N-myristoylation site (G - {EDRKHPFYW}- x(2) – [STAGCN] – {P} where G is the N – myristoylation site), ☐ protein kinase C phosphorylation site ([ST] – x - [RK] where S or T is the phosphorylation site), and ☐ casein kinase II phosphorylation site ([ST] – x(2) – [DE] where S or T is the phosphorylation site). (http://www.expasy.ch/tools/scanprosite/).

The EST sequence of SP02271 contained partial coding region, stop codon with poly(A)$^+$ tail and long 3' UTR, but is lacked of the 5' UTR, the start codon and most of the coding sequence (cds) when compared to the transcripts encoding mannuronan C5-epimerase sequences from *L. digitata*, *E. siliculosus*, *S. japonica* (Figure 4.11). The transcript SP02271 is 1393 bp, after sequencing using universal primers T3 and T7. Nonetheless, the amino acid sequence of SP02271 matches to three known protein profiles from the PROSITE database, including N-myristoylation site (G - {EDRKHPFYW}- x(2) – [STAGCN] – {P} where G is the N – myristoylation site), protein kinase C phosphorylation site ([ST] – x - [RK] where S or T is the phosphorylation site), and casein kinase II phosphorylation site ([ST] – x(2) – [DE] where S or T is the phosphorylation site) (Figure 4.10).



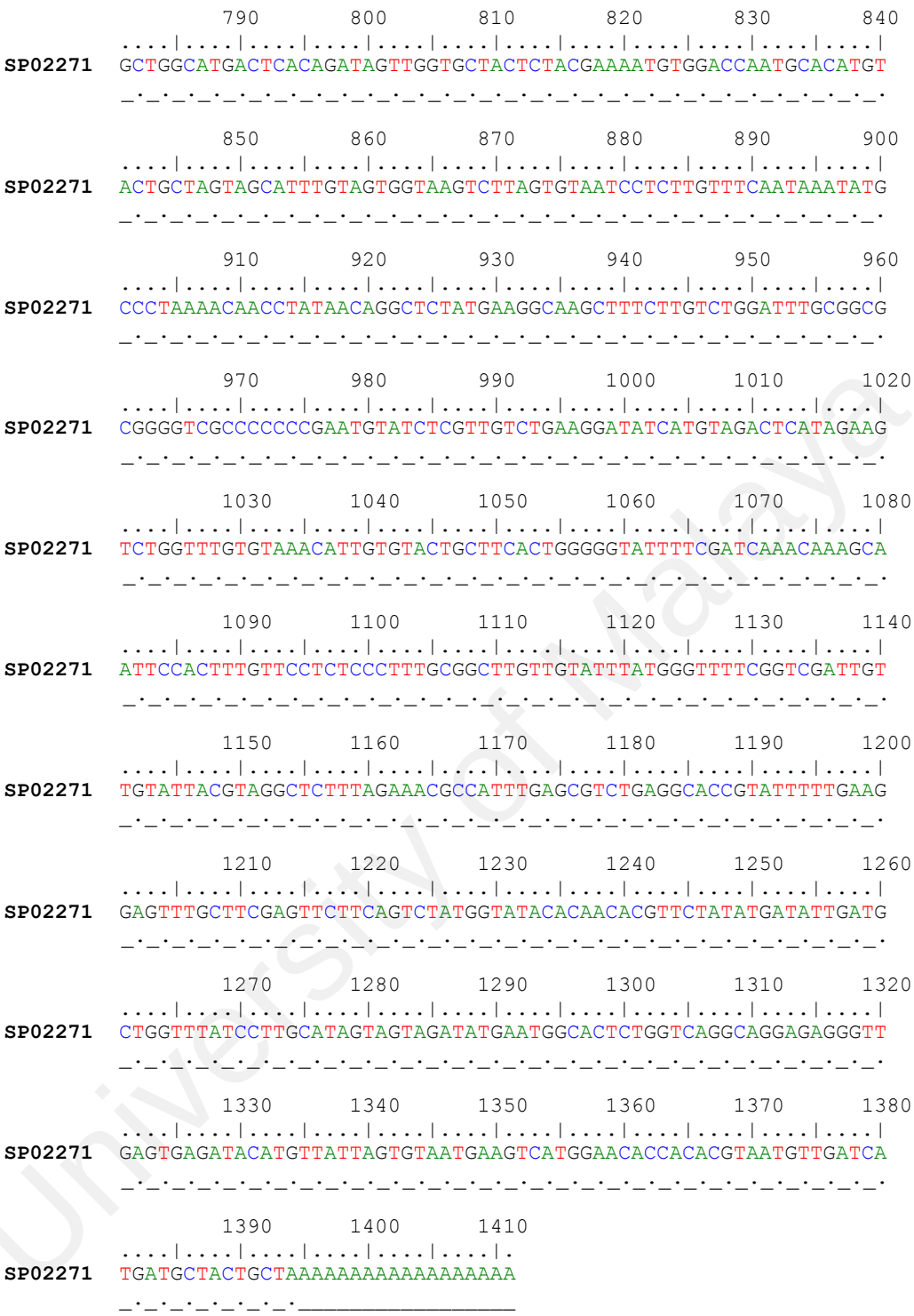
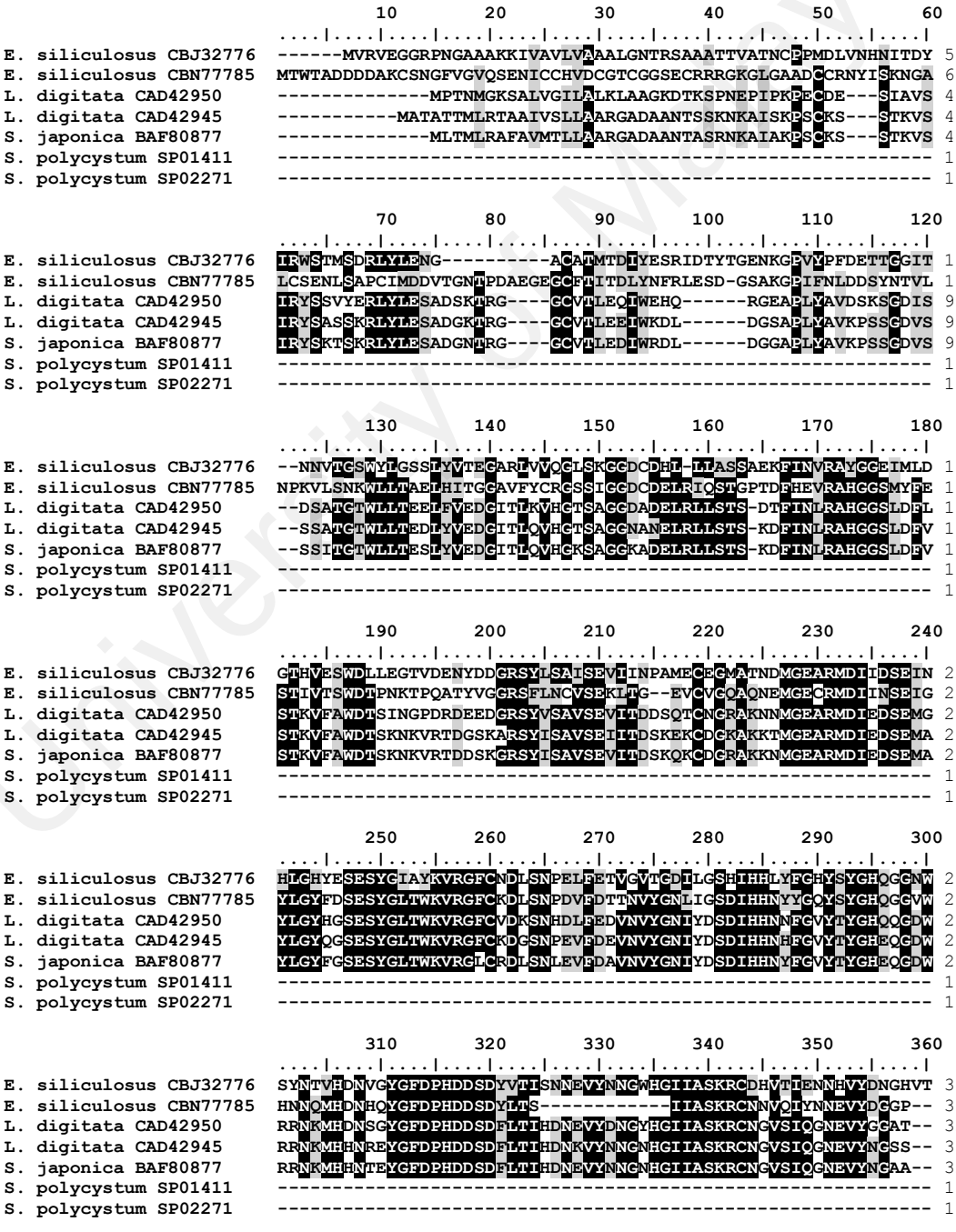


Figure 4.10: Partial transcript sequence encoding mannuronan C5-epimerase (MC5E-2) (SP02271).

The predicted amino acid sequence of the transcript is shown under the nucleotide. _._. indicates the 3' UTR and ___ indicates poly(A)+ tail. For Prosite Scan, ▢ indicates N-myristoylation site (G - {EDRKHPFYW}- x(2) – [STAGCN] – {P} where G is the N – myristoylation site), ▢ protein kinase C phosphorylation site ([ST] – x - [RK] where S or T is the phosphorylation site), and ▢ casein kinase II phosphorylation site ([ST] – x(2) – [DE] where S or T is the phosphorylation site). (http://www.expasy.ch/tools/scanprosite/).

The BLASTP analysis against NCBI nr protein database showed that all the two partial mannuronan C5-epimerase deduced proteins are more closely related to the MC5E from *E. siliculosus* followed by *L. digitata* and then by *S. japonica*. Figure 4.11 shows multiple amino acid sequence alignment of mannuronan C5-epimerase SP01411 and SP02271 from *S. polycystum* and the amino acid sequences from other organisms.



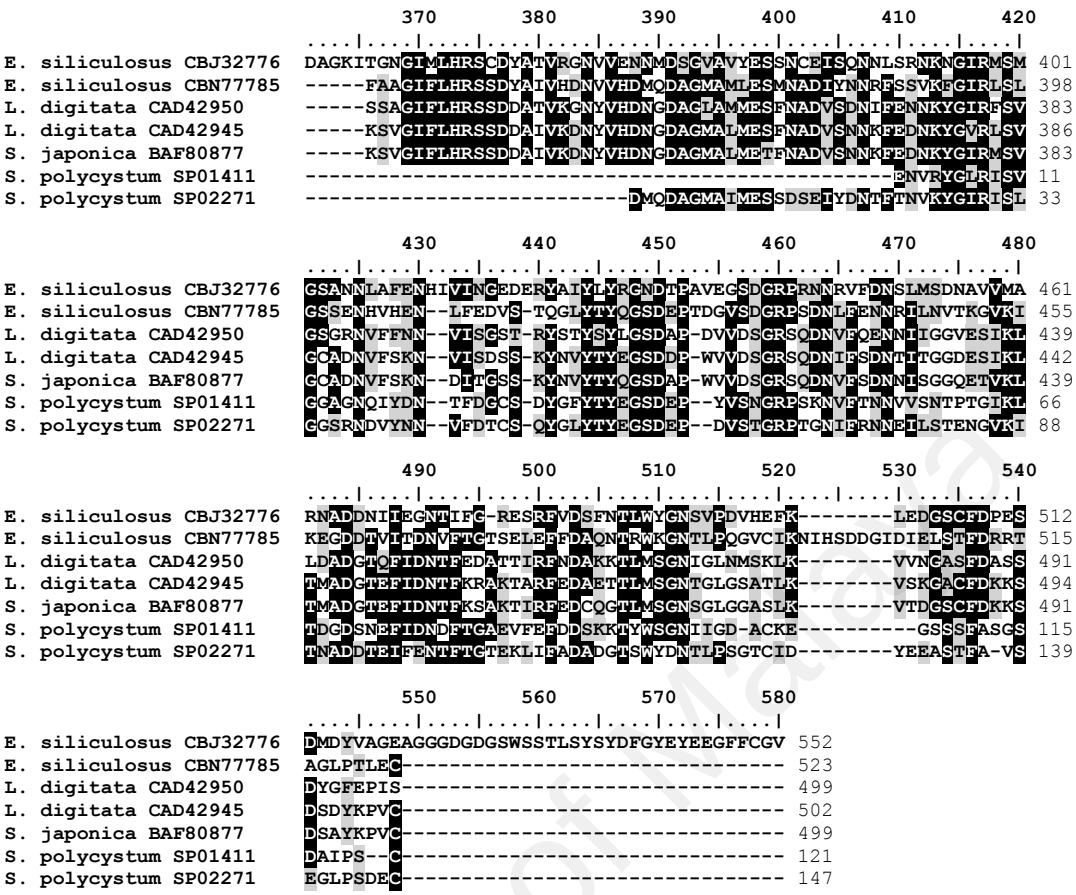


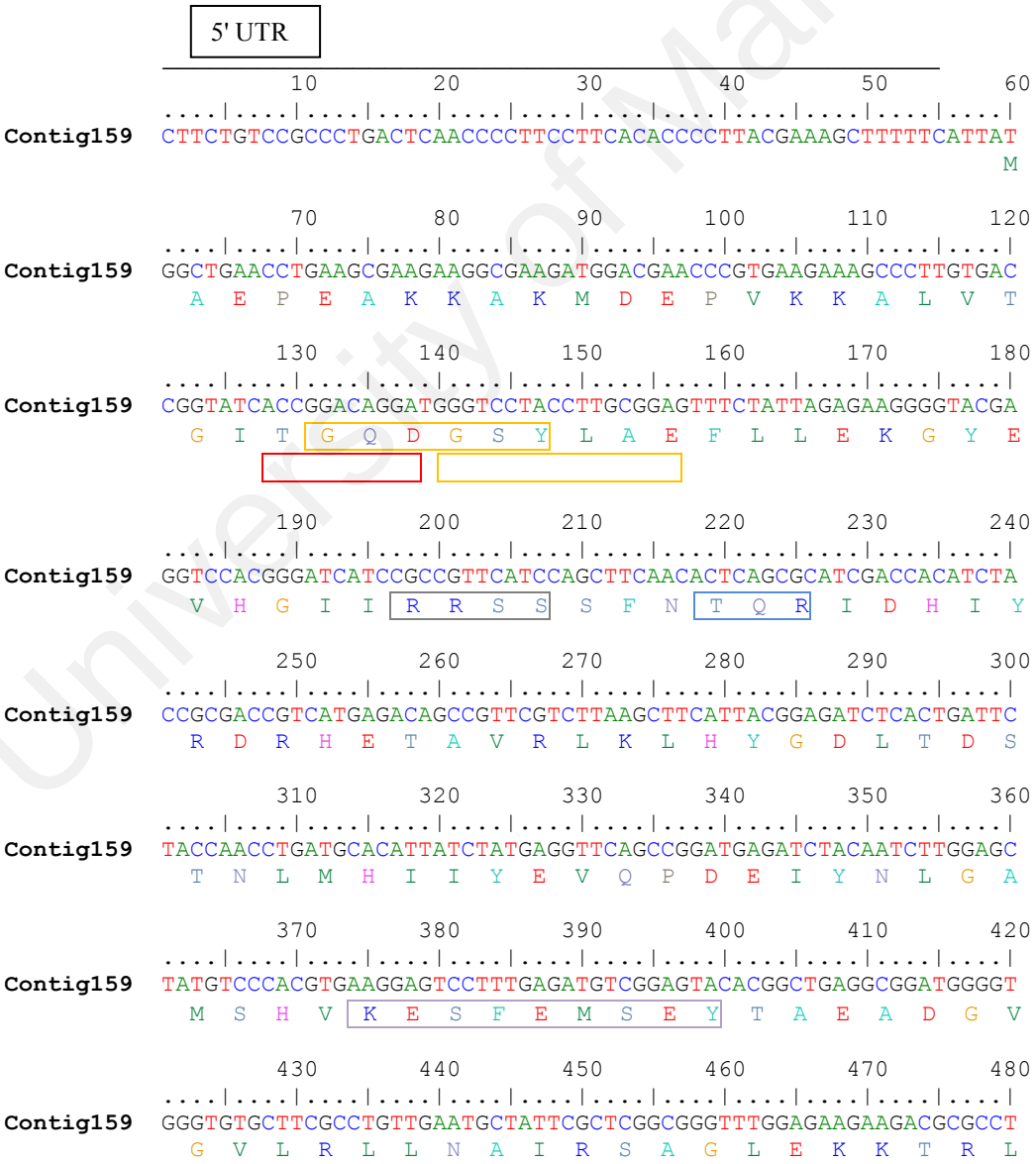
Figure 4.11: Multiple amino acid sequence alignment of partial mannuronan C5-epimerase SP01411 and SP02271 from *S. polycystum* and the complete amino acid sequences from other brown seaweeds.

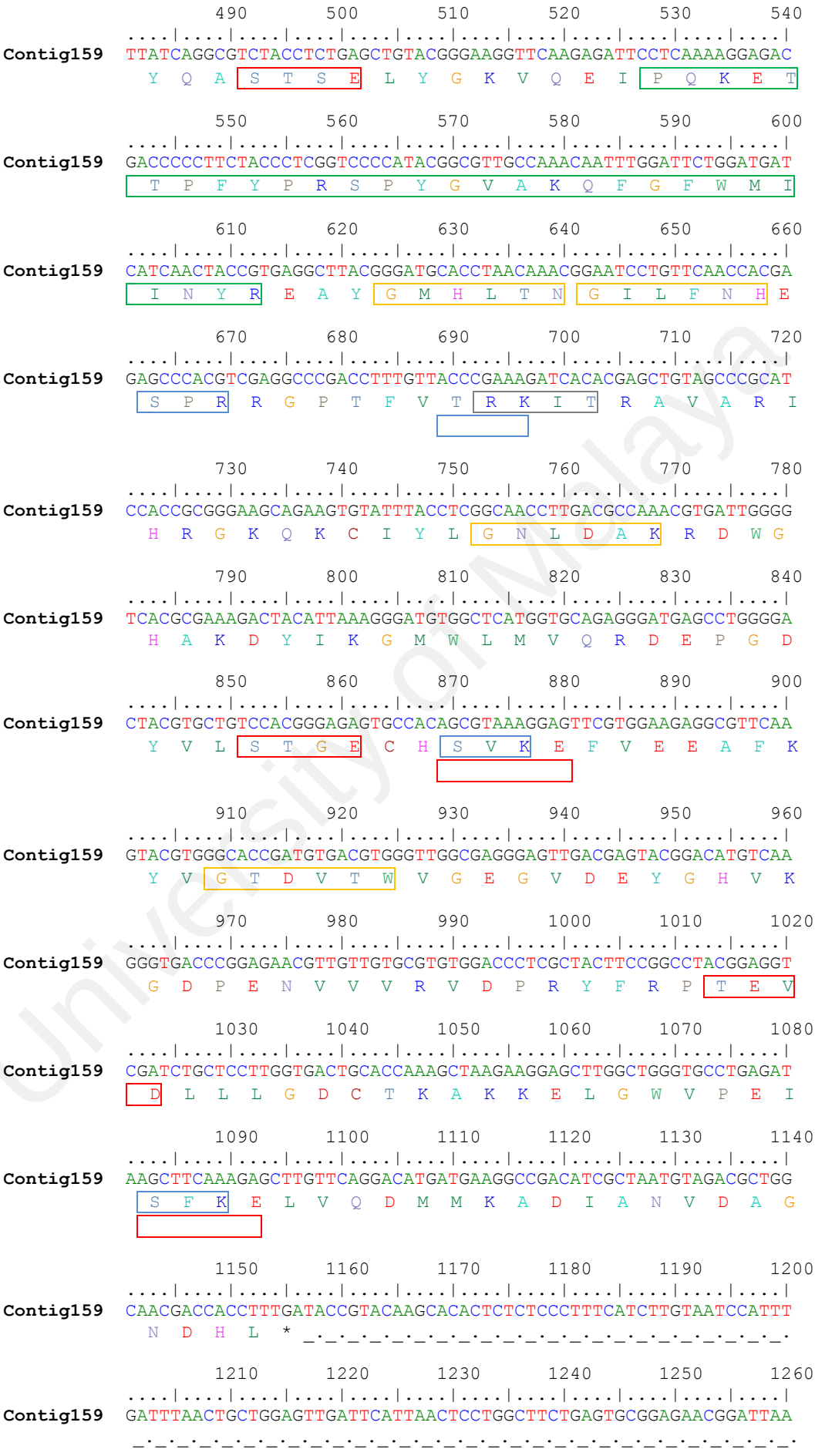
Protein sequences were aligned with ClustalW algorithm using amino acid sequences from *E. siliculosus* (CBJ32776; CBN77785), *L. digitata* (CAD42950; CAD42945), *Saccharina japonica* (BAF80877). The dark-gray shading indicates identical amino-acids; dashes indicate gaps in sequences for the best alignment. *S. polycystum*: *Sargassum polycystum*; *E. siliculosus*: *Ectocarpus siliculosus*; *L. digitata*: *Laminaria digitata*; *S. japonica*: *Saccharina japonica*.

4.6.2 Sequence Analysis of GDP-D-Mannose 4,6 Dehydratase (GM46D)

In this study, two transcripts encoding GDP-D-mannose 4,6 dehydratase (GM46D) were identified among the *S. polycystum* unigenes (Contig159 and Contig254). However, frame shift mutation has occurred in transcript Contig254, thus Contig254 was not used for further analysis. On the other hand, the EST sequence of Contig159 (with a total length of 1753 bp) contained complete coding sequence encoding GM46D

when compared to the transcripts encoding GM46D sequences from *E. siliculosus*, *A. anophagefferens*, *T. pseudonana* and *Blastocystis* sp. (Figure 4.13). Nonetheless, deduced amino acid sequence of Contig159 has matches to six known protein profiles from the PROSITE database, including N-myristoylation site (G - {EDRKHPFYW}-x(2) – [STAGCN] – {P} where G is the N – myristoylation site), protein kinase C phosphorylation site ([ST] – x - [RK] where S or T is the phosphorylation site), and casein kinase II phosphorylation site ([ST] – x(2) – [DE] where S or T is the phosphorylation site) (Figure 4.12).





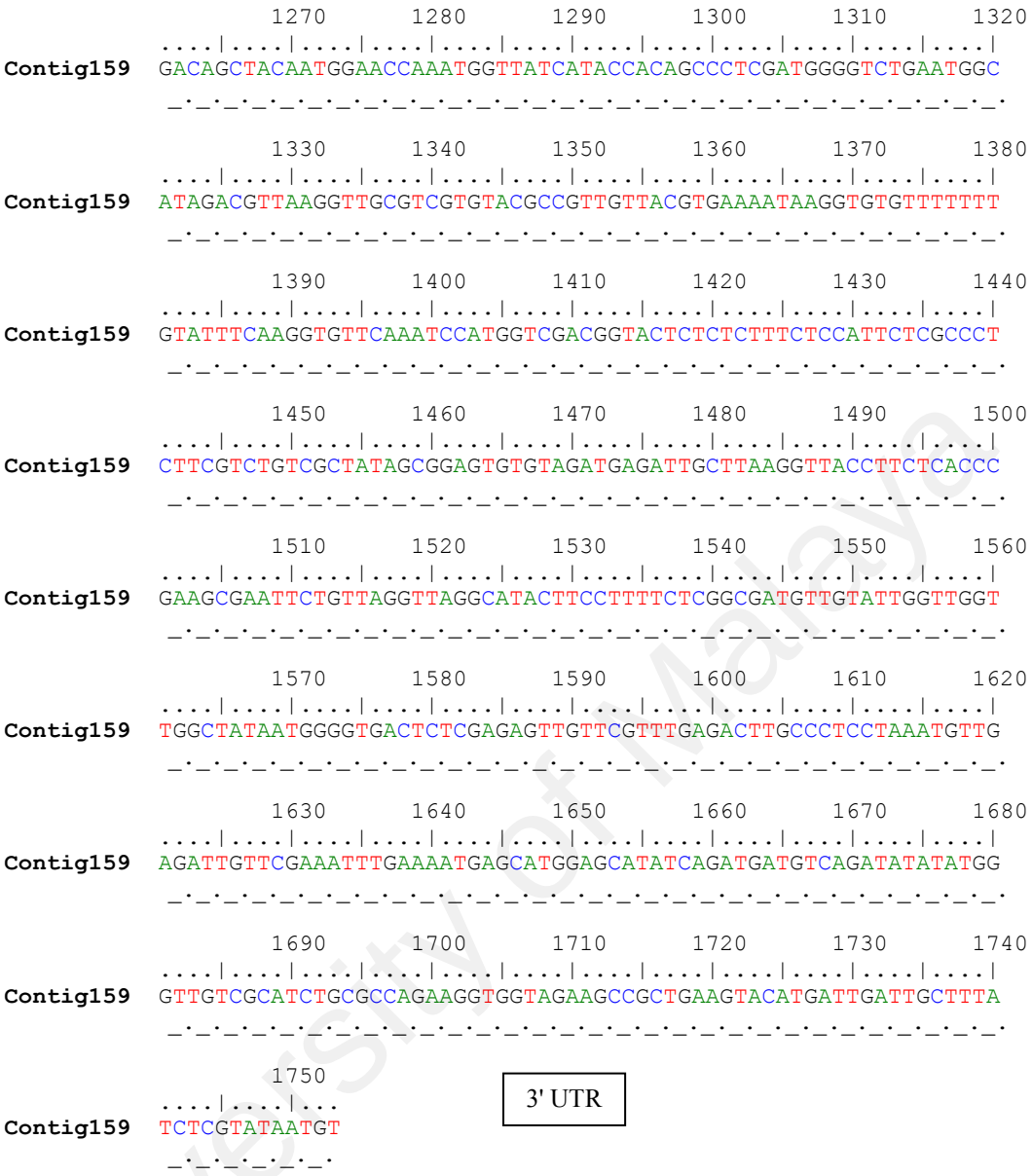
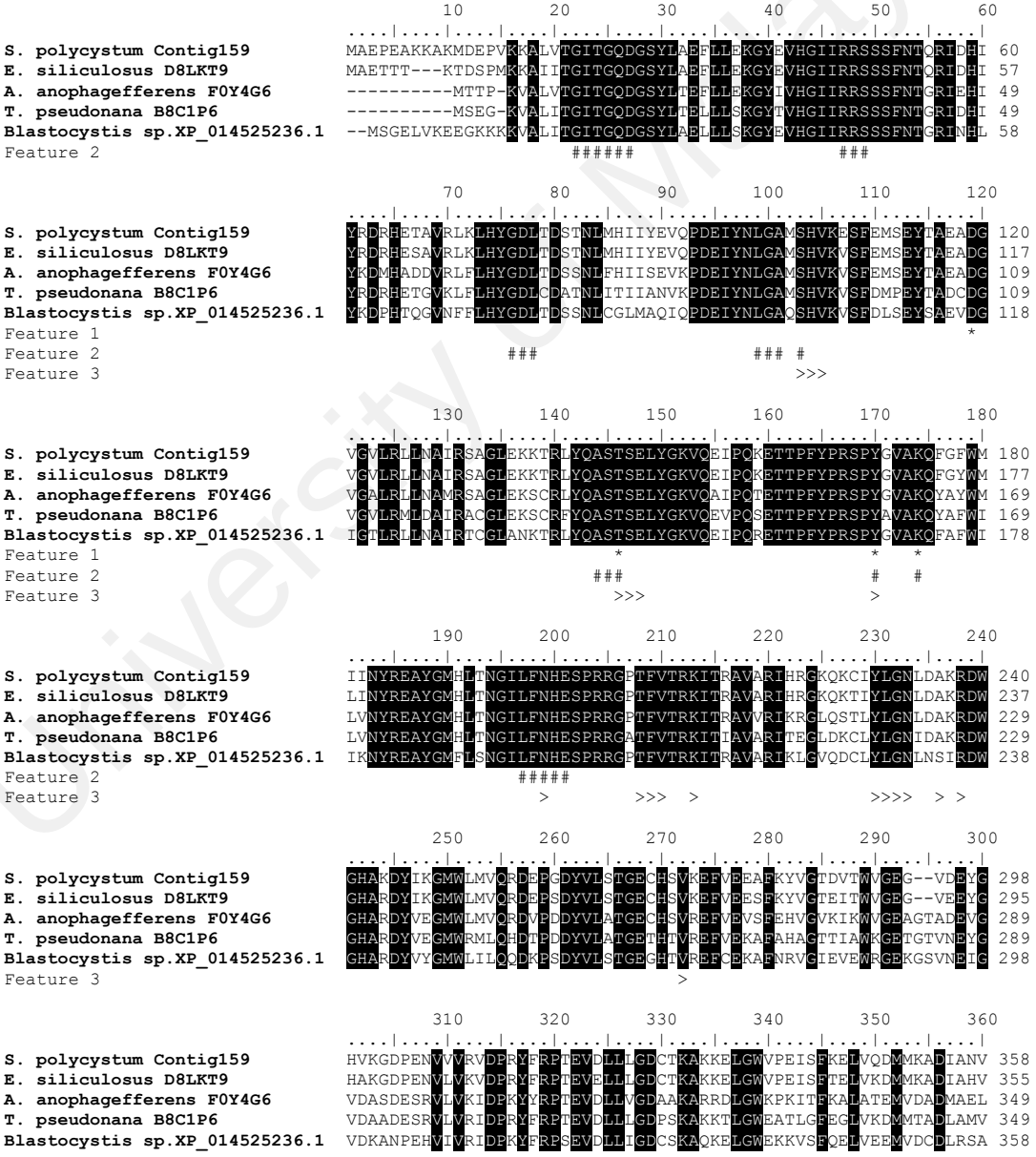


Figure 4.12: Complete coding sequence encoding GDP-D-mannose 4,6 dehydratase (GM46D) (Contig159).

The predicted amino acid sequence of the transcript is shown under the nucleotide. ___ indicates the 5' UTR, _._. indicates the 3' UTR. For Prosite Scan, indicates short-chain dehydrogenase/reductase family signature ([LIVSPADNK]-x(9)-{P}-x(2)-Y-[PSTAGNCV]-[STAGNQCIVM]-[STAGC]-K-{PC}-[SAGFYR]-[LIVMSTAGD]-x-{K}-[LIVMFYW]-{D}-x-{YR}-[LIVMFYWGAPTHQ]-[GSACQRHM] where Y is an active site residue), indicates N-myristoylation site (G - {EDRKHPFYW}- x(2) – [STAGCN] – {P} where G is the N – myristoylation site), cAMP- and cGMP-dependent protein kinase phosphorylation site ([RK](2) – x- [ST] where S or T is the phosphorylation site), protein kinase C phosphorylation site ([ST] – x - [RK] where S or T is the phosphorylation site), casein kinase II phosphorylation site ([ST] – x(2) – [DE] where S or T is the phosphorylation site) and tyrosine kinase phosphorylation

site ([RK]-x(2)-[DE]-x(3)-Yor[RK]-x(3)-[DE]-x(2)-Y where Y is the phosphorylation site) (http://www.expasy.ch/tools/scanprosite/).

Multiple sequence alignment of amino acid sequence encoding GDP-D-mannose 4,6 dehydratase (GM46D) from *S. polycystum* and the amino acid sequences from other organisms was conducted (Figure 4.13). GM46D from *S. polycystum* has 93 % identity to that of *E. siliculosus* (D8LKT9) with a coverage of 97 %.



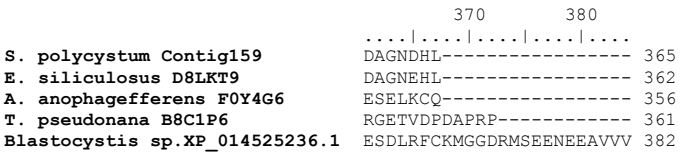


Figure 4.13: Multiple amino acid sequence alignment of GDP-D-mannose 4,6 dehydratase (GM46D) from *S. polycystum* and the amino acid sequences from other organisms.

Protein sequences were aligned with ClustalW algorithm using amino acid sequences from *E. siliculosus* (brown seaweed; D8LKT9), *A. anophagefferens* (marine picoplankton; F0Y4G6), *T. pseudonana* (diatom; B8C1P6), *Blastocystis* sp (unicellular protist; XP_014525236.1). The dark-gray shading indicates identical amino-acids; dashes indicate gaps in sequences for the best alignment; * indicates active site; # indicates NADP-binding site; > indicates substrate binding site. *S. polycystum*: *Sargassum polycystum*; *E. siliculosus*: *Ectocarpus siliculosus*; *A. anophagefferens*: *Aureococcus anophagefferens*; *T. pseudonana*: *Thalassiosira pseudonana*; *Blastocystis* sp: *Blastocystis* species.

Molecular phylogenetic analysis revealed a closer evolutionary relationship between GDP-D-mannose 4,6 dehydratase (GM46D) from *S. polycystum* and that of the Stramenopiles *E. siliculosus* and *Sac. japonica*. On the other hand, GM46D of *S. polycystum* were associated with a major clade of other Stramenopiles (unicellular protist, oomycetes and diatom), Fungi, Amoebozoa, Animalia and Chlorophyta (Figure 4.14). The deduced amino acid showed no signal peptide when analysed with SignalP 4.0.

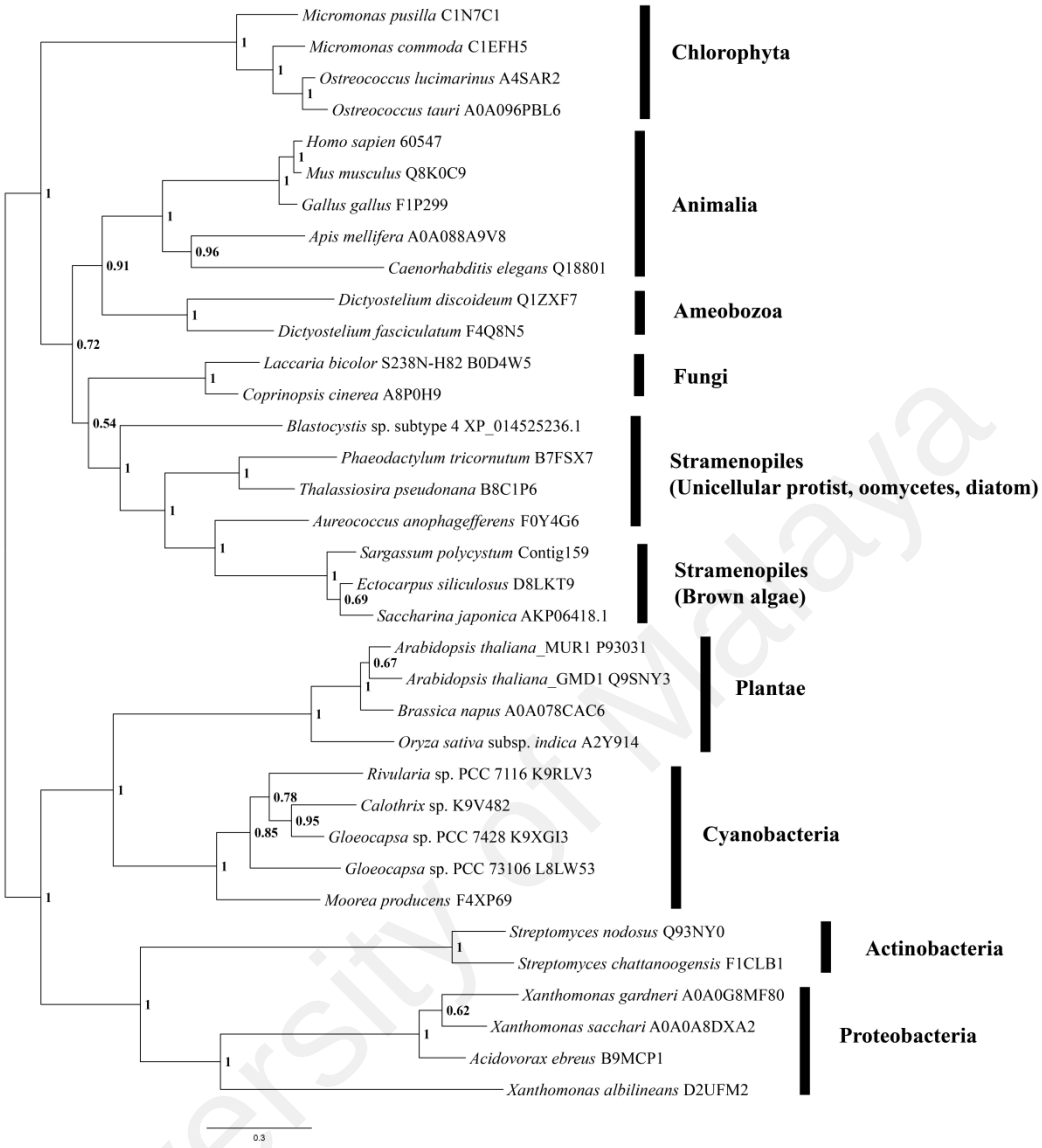


Figure 4.14: The Bayesian phylogenetic tree of GDP-D-mannose 4,6-dehydratase (GM46D).

Numbers at nodes are Bayesian posterior probabilities. The sequences are labeled with species names and their accession number. The unit of branch length is the number of substitutions per site.

Subcellular localization analysis of GM46D by WoLFPSORT showed that it can be targeted to a number of subcellular and extracellular locations (Table 4.8) which include cytoplasmic, endoplasmic reticulum and membrane plasma, whereas subcellular

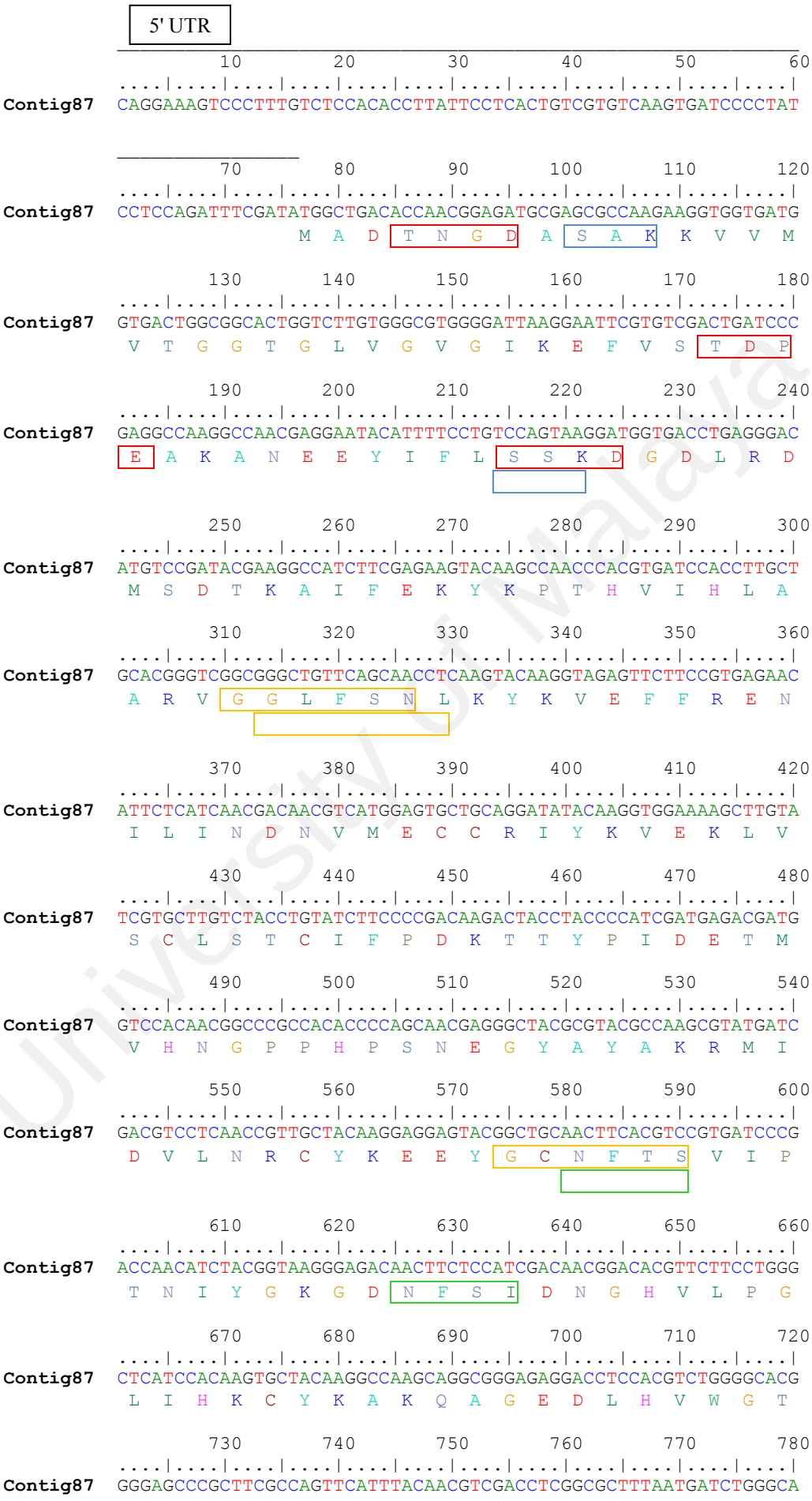
localization of GM46D by HECTAR showed it is a protein with no detectable N-terminal target peptide.

Table 4.8: Cellular localization of GM46D predicted by WoLF PSORT.

Neighbour identity	Localization site	Distance	Identity (%)
At2g27760.1	Cytoplasmic	163.609	14.81
SOD2_BRAJU	Cytoplasmic	229.355	10.41
PMGI_MESCR	Cytoplasmic	231.519	13.77
EF1A_MAIZE	Cytoplasmic	231.962	12.95
EF1A_PEA	Cytoplasmic	232.469	12.30
PMGI_TOBAC	Cytoplasmic	232.702	14.31
PMGI_RICCO	Cytoplasmic	238.352	15.11
NRL3_ARATH	Membrane plasma	238.578	12.30
At5g13930.1	Endoplasmic reticulum	239.188	12.66
ACT8_ARATH	Cytoplasmic	240.094	10.34
ACT12_ARATH	Cytoplasmic	241.537	10.32

4.6.3 Sequence Analysis of GDP-4-Keto-6-D-Mannose Epimerase/Reductase (GFS)

A transcript encoding GDP-4-keto-6-D-mannose epimerase/reductase (GFS) was identified among the *S. polycystum* unigenes (Contig87). Contig87 (with a total length of 1742 bp) contained complete coding sequence encoding GFS when compared to the transcripts encoding GFS from *E. siliculosus*, *A. aphagefferens*, *N. gaditana* and *L. digitata* (Figure 4.16). The deduced amino acid sequnce of Contig87 has matches to four known protein profiles from the PROSITE database including N-myristoylation site (G - {EDRKHPFYW}- x(2) – [STAGCN] – {P} where G is the N – myristoylation site), protein kinase C phosphorylation site ([ST] – x - [RK] where S or T is the phosphorylation site), casein kinase II phosphorylation site ([ST] – x(2) – [DE] where S or T is the phosphorylation site) and N-glycosylation site (N – {P} – [ST] – {P} where N is the glycosylation site) (http://www.expasy.ch/tools/scanprosite/). (Figure 4.15).



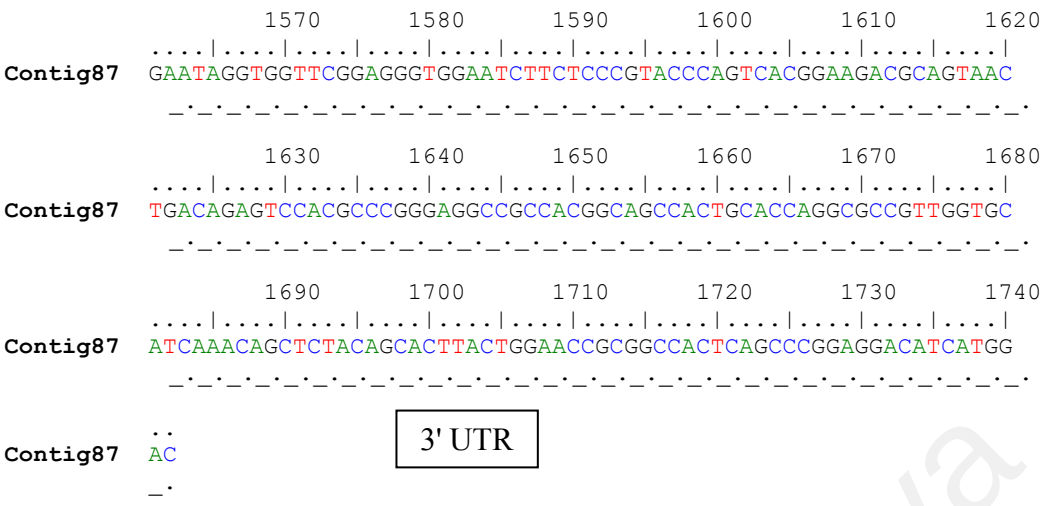
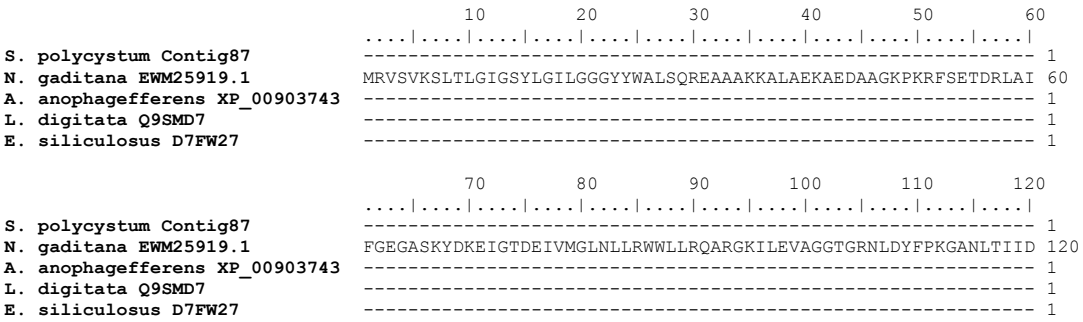


Figure 4.15: Full length transcript sequence of GDP-4-keto-6-deoxy-D-mannose-3,5-epimerase/reductase (GFS) (Contig87).

The predicted amino acid sequence of the transcript is shown under the nucleotide. ___ indicates the 5' UTR, _._. indicates the 3' UTR. For Prosite Scan, indicates N-myristoylation site (G - {EDRKHPFYW}- x(2) – [STAGCN] – {P} where G is the N – myristoylation site), protein kinase C phosphorylation site ([ST] – x - [RK] where S or T is the phosphorylation site), casein kinase II phosphorylation site ([ST] – x(2) – [DE] where S or T is the phosphorylation site) and N-glycosylation site (N – {P} – [ST] – {P} where N is the glycosylation site) (http://www.expasy.ch/tools/scanprosite/).

Multiple sequence alignment of amino acid sequence encoding GDP-4-keto-6-D-mannose-epimerase-4-reductase (GFS) from *S. polycystum* (Contig87) and the amino acid sequences from other organisms was conducted (Figure 4.16). GFS from *S. polycystum* has 94 % identity to that of *E. siliculosus* (D7FW27) with a coverage of 95 %.



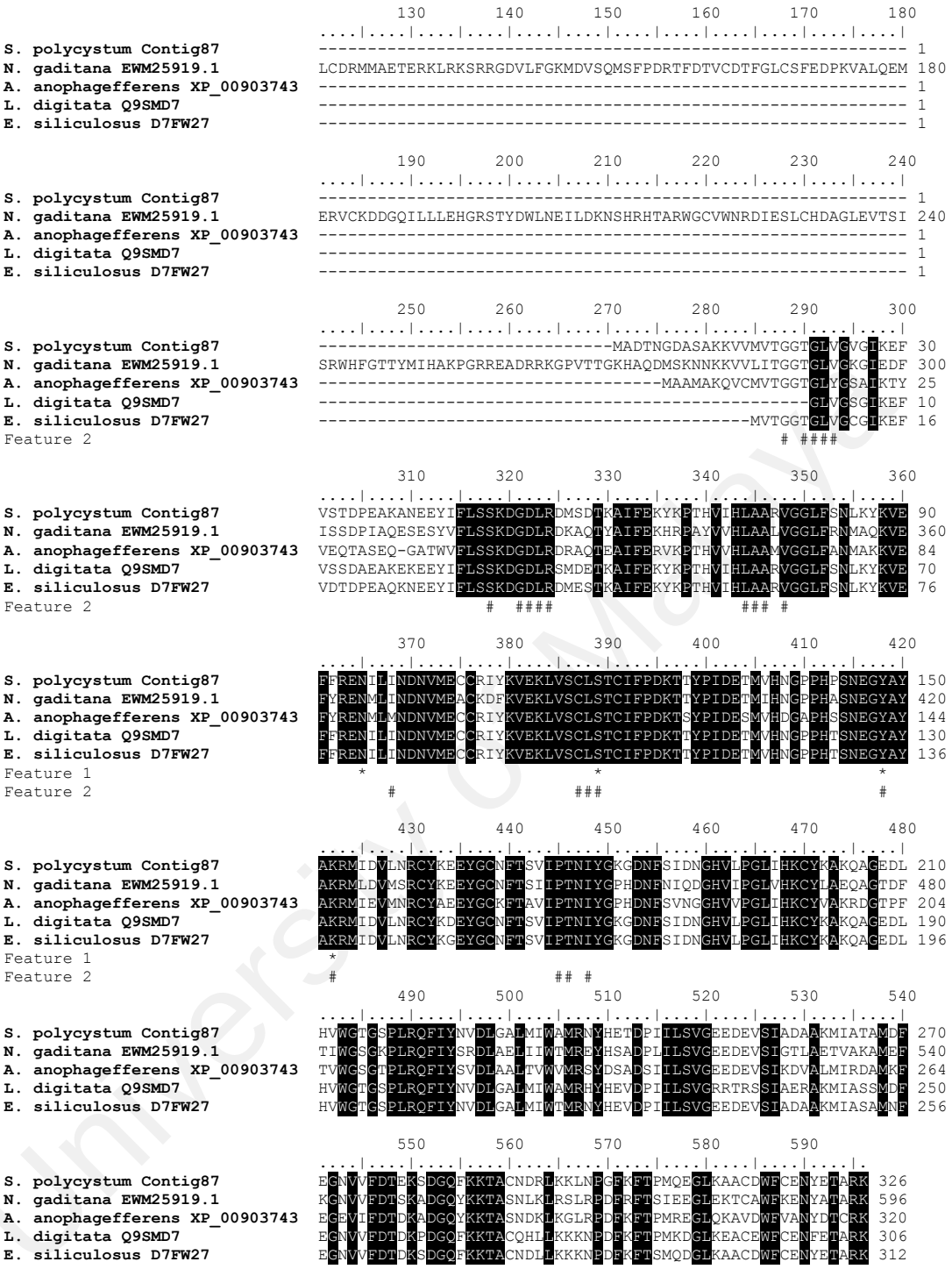


Figure 4.16: Multiple amino acid sequence alignment of GDP-4-keto-6-D-mannose-epimerase-4-reductase (GFS) of *S. polycystum* (Contig87) and the amino acid sequences from other organisms.

Protein sequences were aligned with ClustalW algorithm using amino acid sequences from *E. siliculosus* (brown macroalga; D7FW27), *A. anophagefferens* (marine picoplankton; XP_00903743), *N. gaditana* (marine microalga; EWM25919.1), *L. digitata* (brown macroalga; Q9SMD7). The dark-gray shading indicates identical amino-acids; dashes indicate gaps in sequences for the best alignment. * indicate active site; # indicates NADP binding site; > indicates putative substrate binding site. *S. polycystum*: *Sargassum polycystum*; *E.*

***siliculosus*: *Ectocarpus siliculosus*; *A. anophagefferens*: *Aureococcus anophagefferens*; *N. gaditana*: *Nannochloropsis gaditana*; *L. digitata*: *Laminaria digitata*.**

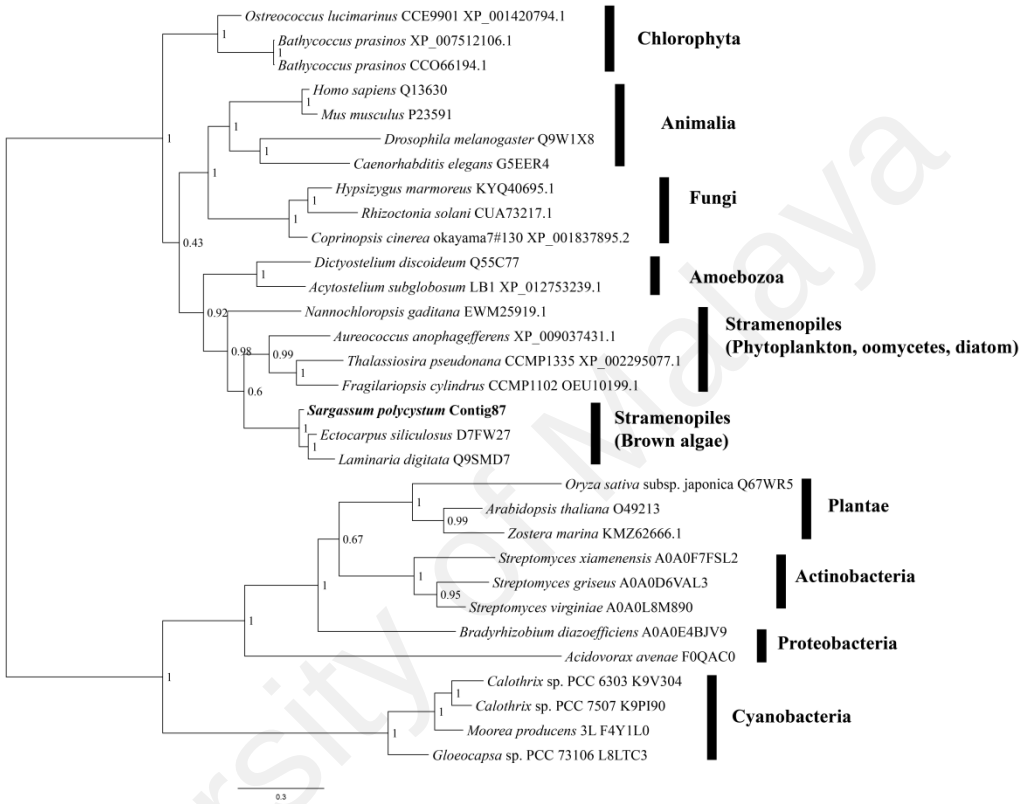


Figure 4.17: The Bayesian phylogenetic tree of GDP-4-keto-6-D-mannose-epimerase-4-reductase (GFS).

Numbers at nodes are Bayesian posterior probabilities. The sequences are labeled with species names and their accession number. The unit of branch length is the number of substitutions per site.

Molecular phylogenetic analysis revealed a closer evolutionary relationship between GDP-4-keto-6-D-mannose-epimerase-4-reductase (GFS) from *S. polycystum* and that of the Stramenopiles *E. siliculosus* and *L. digitata*. On the other hand, GM46D of *S. polycystum* were associated with a major clade of other Stramenopiles (phytoplankton,

oomycetes and diatom), Fungi, Amoebozoa, Animalia and Chlorophyta (Figure 4.17). The deduced amino acid showed no signal peptide when analysed with SignalP 4.0.

Subcellular localization analysis of GFS by WoLFPSORT showed that it is most probably localized in cytoplasm (Table 4.9), whereas subcellular localization by HECTAR showed it is a protein with no detectable N-terminal target peptide.

Table 4.9: Cellular localization of GFS predicted by WoLF PSORT.

Neighbour identity	Localization site	Distance	Identity (%)
GLN2_MAIZE	Cytoplasmic	175.20	12.22
EF1A_LYCES	Cytoplasmic	192.94	14.73
EF11_DAUCA	Cytoplasmic	196.19	14.48
EF1A_ARATH	Cytoplasmic	201.57	14.70
EF1A_MANES	Cytoplasmic	207.75	14.25
At3g02230.1	Golgi apparatus	210.256	12.32
ADH1_PENAM	Cytoplasmic	210.374	12.93
ACT5_ARATH	Cytoplasmic	212.155	14.29
At3g53140.1	Cytoplasmic	219.069	10.86
ADH1_ORYSA	Cytoplasmic	219.419	12.70
ADH_FRAAN	Cytoplasmic	221.446	14.21
ADH1_HORVU	Cytoplasmic	221.728	12.14
ADH1_MAIZE	Cytoplasmic	223.913	12.40

4.7 Selection of Endogenous Genes

The gene expression stability of four endogenous genes (AT, RPL3a, RPS15a and TEF1a) was performed using geNorm version 3.5 software. RPS15A was exluded as the least stable genes (the highest M values of 0.357). The reference genes, RPL3A (M value = 0.192), TEF1A (M value = 0.192) and AT (M value = 0.211) (Figure 4.16), had

the lowest M values and were identified as the most stable reference genes for this set of samples.

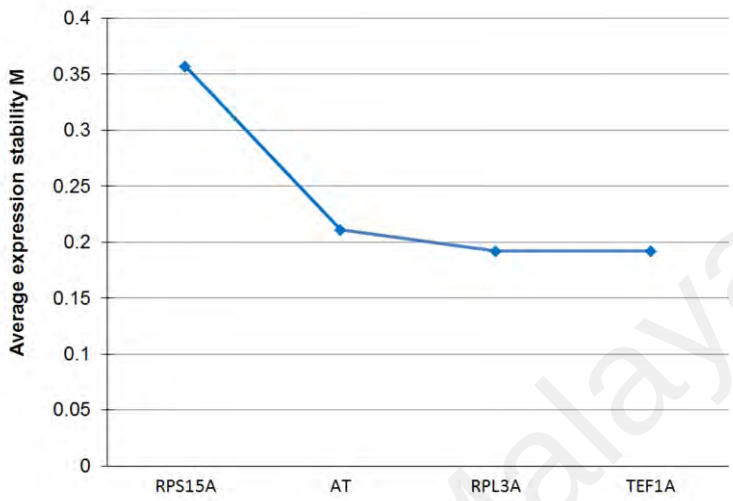


Figure 4.18: The values of stability of the housekeeping genes, and ranking obtained from geNorm pairwise analysis.

The housekeeping genes, RPL3A (M=0.192), TEF1A (M=0.192), AT (M=0.211) had the lowest M values and were identified as the most stable housekeeping genes for this set of samples.

Dissociation curves confirm the amplification specificity, and show that all the primer pairs for transcripts encoding MC5E, GM46D (Contig159) and GFS (Contig87) and housekeeping genes, respectively, exhibited a single peak with no visible primer-dimer formation (Appendix H). The efficiency values of selected primers ranged from 91.4 % for Alpha Tubulin (AT) to 97.1 % for MC5E-1 (Appendix G), implying the RT-qPCR in the study was efficient and specific to *S. polycystum* candidate transcripts amplification.

4.8 Expression Profiles of MC5E in *Sargassum polycystum* in Rainy and Dry Seasons

In this study, the transcript abundance of MC5E-1 (SP01411) in rainy season increased to 2.6-fold compared to that in the dry season (as the control, 1 fold) (Figure 4.19A). In contrast, the transcript abundance of MC5E-2 (SP02271) in rainy season reduced to 32 % of the transcript abundance during dry season (Figure 4.19B). Ct-range values for transcripts MC5E-1 and the reference genes are from 14.32 to 24.75, and Ct-range values for trancripts MC5E-2 and reference genes are from 14.14 to 23.72.

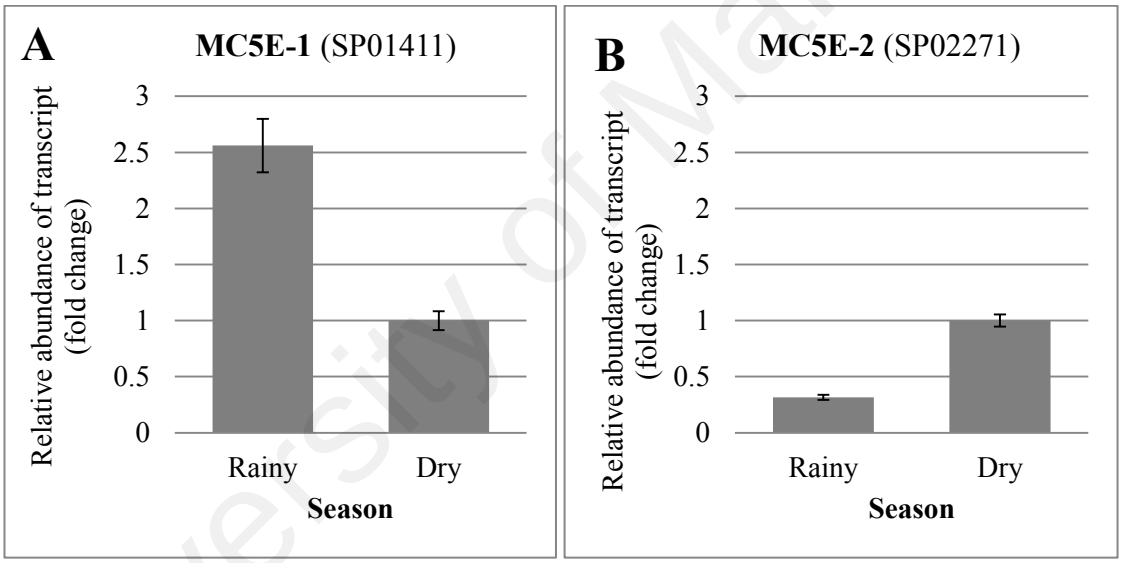


Figure 4.19: Expression levels of cDNAs encoding MC5E of *S. polycystum*.

(A) MC5E-1 (SP01411); (B) MC5E-2 (SP02271). After normalization to the transcript abundance of TEF1A, AT and RPL3A, the transcript abundance was expressed as relative fold-change to that of the controls. The standard errors were calculated from the results of three replicates.

4.9 Expression Profiles of GM46D (Contig159) and GFS (Contig87) in *Sargassum polycystum* in Rainy and Dry Seasons

The transcript abundance of GM46D-Contig159 reduced to 71 % of the transcript abundance during dry season (Figure 4.20A) in this study. In contrast, the transcript abundance of GFS-Contig87 was increased to 1.2-fold during rainy season compared to that in the dry season (Figure 4.20B). Ct-range values for transcripts GM46D-Contig159 and the reference genes are from 14.39 to 19.93, and Ct-range values for trancripts GFS-Contig87 and reference genes are from 14.39 to 19.06.

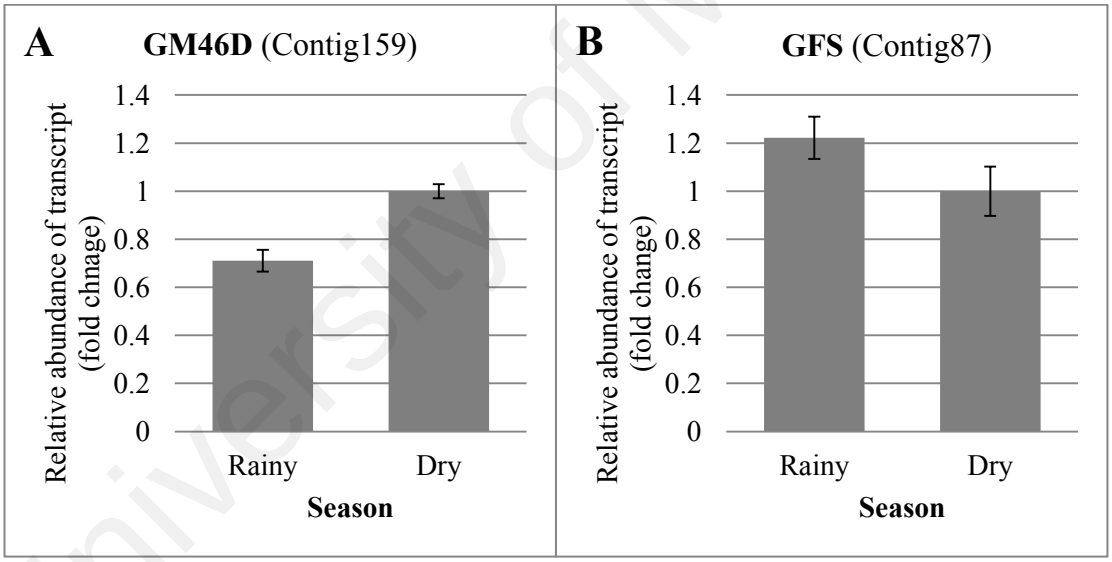


Figure 4.20: Gene expression profiles of cDNAs encoding GM46D and GFS in *S. polycystum*.

(A) GM46D (Contig159); (B) GFS (Contig87). After normalization to the transcript abundance of TEF1A, AT and RPL3A, the transcript abundance was expressed as relative fold-change to that of the controls. The standard errors were calculated from the results of three replicates.

CHAPTER 5: DISCUSSION

5.1 Optimised RNA Isolation Protocol for *S. polycystum*

The RNA isolation from plants as well as seaweeds can be a major problem due to the large amounts of polysaccharides and polyphenolic compounds during the developmental stages and/or in response to environmental conditions (Tonon *et al*., 2008; Caffall & Mohnen, 2009). These polysaccharides and polyphenolic compounds often coprecipitate with nucleic acids during the extraction process, thus affecting both the quality and quantity of isolated nucleic acids (Wang *et al*., 2008). Seaweed tissues have high levels of polysaccharides and secondary metabolites that are released after disruption of cells embedded in viscous polysaccharides (Ho *et al*., 1996). *S. polycystum* has a high content of polysaccharides, particularly alginate, where the yield of alginate obtained ranged from 17.12 to 39.8 % (Saraswathi *et al*., 2003; Goh, 2008).

Guanidinium thiocyanate and guanidinium chloride are effective protein denaturants and inhibitor of RNase (Chomczynski & Sacchi, 1987). However, these strong protein denaturants may not be suitable for tissues that are rich in polysaccharides or phenolic compounds (Wilkins & Smart, 1996). The presence of high amounts of polyphenols and secondary metabolites has been shown to interfere with RNA extraction in the presence of guanidinium salts (Bugos *et al*., 1995). Guanidinium salts are ineffective in dissociating RNA from nonprotein complexes, and RNA may be lost along with the complex during the isolation procedure (Liu *et al*., 1998; Bugos *et al*., 1995). This may have happened to the total RNA extracted from *S. polycystum* using method 1 whereby the RNA was sticky, brown in color, and difficult to completely dissolve in water even after incubation at 60 °C. Apart from that, low A_{260}/A_{280} ratio (1.175 ± 0.064) and low A_{260}/A_{230} ratio (0.535 ± 0.064) obtained from method 1 suggested that guanidinium thiocyanate is not suitable for removing protein and polysaccharides from *S.*

polycystum.

Method 2 used lithium chloride in the lysis buffer. The seaweed tissues could be soften by the aid of LiCl and the nucleic acids would be released from the loosened cell membrane and as well as cell wall (Hong *et al*., 1997). The lysis buffer which containing LiCl managed to eradicate polysaccharide contamination from nucleic acids in *Gracilaria changii* (Chan *et al*., 2004), marine macroalgae (Kim *et al*., 1997) and *Porphyra* (Hong *et al*., 1997), but it did not work on *S. polycystum*, as judged from the gelatinous pellet obtained in this study. The presence of polyphenolic compounds made the RNA pellet difficult to dissolve and the presence of polysaccharides made the RNA pellet solution viscous (Accerbi *et al*., 2010).

The PVPP in method 3 is known to form complex with polyphenolics through the hydrogen bond and thus helps in the removal of polyphenolics from the homogenate. However, this method produced a dark brown pellet which was difficult to dissolve. The colored pellets (brown pellets) could be a result of the oxidation and decomposition of phenolic compounds (Accerbi *et al*., 2010). This method also suggested that PVPP is not suitable for removing protein and polysaccharides from *S. polycystum*.

Method 4 was not efficient, probably due to the low concentrations of DTT and NaCl which resulted in a low-quality RNA from *S. polycystum*. Method 5 is an improved method from Wong *et al*. (2007) (Method 4). The used of CTAB in this method was as the cell-disrupting agent. Yao *et al*. (2009) have used the CTAB method to isolate total RNA from *Laminaria* (*Saccharina*) *japonica*. Whereas Apt *et al*. (1995) have used the CTAB method to isolate total RNA from the brown alga *Macrocystis pyrifera*. However, Method 5 has some advantages over their methods. Firstly, the use of chaotropic agents such as phenol and β-mercaptoethanol, which are very toxic and

carcinogenic, were eliminated. In addition, phenol extraction during the extraction process can damage poly(A)$^{+}$ RNA (Azevedo *et al*., 2003). Even though β-mercaptoethanol and phenol were not used in the extraction buffer, the integrity of the RNA was still maintained. Secondly, high concentrations of DTT (50 mM), instead of 20 mM DTT, were used to inhibit RNase activity during homogenization step. DTT as the reducing agent, where a high concentration of DTT possesses protein-denaturing activity which effectively inhibits RNase activity by disrupting disulfate bond formation and also prevents oxidation of nucleic acids by phenolic compounds and other unwanted reactions (Pearson *et al*., 2006). to prevent any oxidation Thirdly, the inclusion of absolute ethanol, before the second CI extraction, is critical in obtaining intact RNA. When absolute ethanol was not added during the process of extraction, the integrity of isolated RNA was extremely low. Yao *et al*. (2009) used one quarter absolute ethanol and one ninth volume potassium acetate (pH 4.8) to precipitate polysaccharides into salts. But method 5 used one third volume of ethanol alone, which is sufficient to precipitate the polysaccharides. Furthermore, potassium acetate is not suitable for removing polysaccharides from *G. changii* (Chan *et al*., 2004). Next, a final concentration of 2 M LiCl was used to precipitate the RNA selectively. A higher LiCl concentration increased the amount of RNA impurities such polysaccharides, despite of a higher RNA yield. Results of two hours precipitation by using LiCl was adequatly good compared to the overnight precipitation by using LiCl as suggested by Wong *et al*. (2007) and Apt *et al*. (1995). Lastly, the duration of RNA precipitation was shortened from overnight to 2 h, but at a lower temperature (-80 °C).

In this study, Method 5 was also suitable used for other brown seaweeds (Sim *et al*., 2013). The RNA yield of *S. polycystum* and *S. binderi* were 43.52 ± 0.716 and 25.99 ± 2.467 μg g^{-1} fresh weight, respectively. Even though they are taxonomically very close,

and considered to have very similar cytoplasmic and cell compositions, the differences in the yield of RNA in this study maybe due to alginate and/or phlorotannin content in these two brown seaweeds are varied and also probably due to different collection time. Phlorotannins which only found in brown algae are influenced by surrounding environment such as ultraviolet radiation, nutrient availabilty and light. Firdaus *et al*. (2015) reported that phlorotannin content found in *Sargassum* species can be varied from 13 – 39 mg g^{-1}.

5.2 Generation of ESTs from *S. polycystum*

The RNA obtained from this method was suitable for the isolation of mRNA. The isolation of mRNA from total RNA is a critical step prior to a cDNA library construction. Sufficient amount of good quality mRNA was required, in order to construct a good cDNA library. In eukaryotes, most mRNAs are polyadenylated and these 3'-tails of about 200 adenine residues each are useful for the isolation of eukaryotic mRNA. Under high-salt conditions, these tails are able to form stable hybrids with oligo(dT). This method is efficient for removing the non-mRNA RNAs. However, this method has a drawback which it does not favour for mRNAs that lack poly$(A)^+$ tail, such as deadenylated mRNAs whose poly$(A)^+$ tails have been removed during mRNA turnover or normal histone mRNAs (Rio *et al*., 2010). Bacterial which are prokaryotes do not have mRNA with poly$(A)^+$ tail, thus the isolation of mRNA using oligo(dT) was able to exclude mRNA from possible contaminating prokaryotes. These precautions were taken mainly to ensure that the isolated mRNA was from *S. polycystum*. Oligo(dT) that binds to the poly$(A)^+$ tail is used to recover the mRNA, which was provided in the μMACS mRNA isolation kit. The percentage of mRNA

recovery was satisfactory (~ 1 – 2 %). The titer of primary library in this study was 1 x 10^6 pfu/mL and the titer of the amplified library was 4.2 x 10^{10} pfu/mL. A cDNA library constructed in phage vector contain 10^6 to 10^7 independent recombinants is sufficient to contain even the rarest mRNA (Seidman, 2006). This indicated that the *S. polycystum* library in this study might sufficient to identify novel transcripts and thus can provide useful information.

A total of 2577 (99 %) EST sequences were selected for further analysis after bad quality and short sequences (less than 100 bp) were discarded. Short sequences that were less than 100 bp could be aligned to subjects in the database easily compared to long sequences, thus, increasing the possibility of mismatches between query sequences to subjects.

This study lays the foundation for a broader aim to generate a comprehensive list of coding sequences. Samples for this study were collected from different individuals at different months at Teluk Kemang, Port Dickson. Al-Swailem *et al*. (2010) suggested that inclusion of a more diverse sampling would yield a better coverage of their camel transcriptome study. Therefore, inclusion different individual of *S. polycystum* samples might yield a better coverage in this study. This current study has generated 2577 ESTs for an initial phase study to identify *S. polycystum* genes.

5.3 Functional Annotation of *S. polycystum* ESTs

The sequence analysis was performed using Blast2GO by comparing the assembled ESTs (unigenes) from *S. polycystum* against Swiss-Prot protein database and NCBI non-redundant protein (nr) database with an e-value cut-off of 1e-5, respectively. Most

of the ESTs in this study showed high homology to Ochrophyta, and 79.8 % was assigned to *E. siliculosus* as it is the only species which has been fully sequenced and annotated to date (Cock *et al*., 2010).

Of 1724 unigenes, 1142 (66.2 %) unigenes were matched to NCBI non-redundant protein database. Of 1142 unigenes, contaminants from diatom, bacteria and oomycetes among the top matches of BLASTX were only 1.3 %, 1.1 % and 1 %, respectively. These contaminants are quite small (1 - 1.3 %) since the number of ESTs sequenced was quite small and a non-normalized library was used. It will be more obvious if we increase the depth of sequencing.

Of 1724 unigenes, 926 unigenes (~53.7 %) has no functional information available when Swiss-Prot protein database was used, while 582 (~33.8 %) has no functional information available when nr protein database was used. However, there were many unigenes (34 %) that were assigned as ‗hypothetical protein', ‗expressed unknown protein', ‗conserved unknown protein' and ‗predicted protein' based on sequence similarity to entries in the NCBI nr protein database (BLASTX; E-value $\leq 10^{-5}$). All these unigenes encode proteins with an unknown/novel functions; these proteins are highly diverged, or alternatively, are fragments of protein-encoding unigenes that are not conserved (such as 5'- or 3'- UTR) (Chan *et al*., 2012b). These unigenes are considered as novel genes that may be potentially important for new gene discovery. Brown algal transcripts usually have long 3' UTR (Crépineau *et al*., 2000; Cock *et al*., 2010; Pearson *et al*., 2010; Liu *et al*., 2014). Most of the time, non-overlapping fragments from the same transcript are not able to be assembled into one contig, thus leading to redundancies. In addition, small contigs might belong to different UTRs of the same gene locus (Heinrich *et al*., 2012). Pearson *et al*. (2010) concluded that annotation failure is mainly due to a lack of sufficient coding sequences where long 3'

UTR sequences hindered transcription occured. In fact, low overall annotation rates was shown by available brown algae EST libraries (Roeder *et al*., 2005; Wong *et al*., 2007; Pearson *et al*., 2010). Wang *et al*. (2012) have shown that the expression levels and the sequence lengths were positively relative to the annotation rate in *S. japonica*. Low overall annotation rates also demonstrate that the sequenced genome data of chromalveolates and *E. siliculosus* in the public databases are not sufficient to analyse the sequences of *S. polycystum*. To date, only the genomes of two brown algae that are *E. siliculosus* (Cock *et al*., 2010) and *S. japonica* (Ye *et al*., 2015) and a few chromalveolates that are *T. oceanica* (Lommer *et al*., 2010), *T. pseudonana* (Armbrust *et al*., 2004) and *P. tricornutum* (Bowler *et al*., 2008) have been fully sequenced. Manual curation for *E. siliculosus* genome was still in progress. However, once the complete genome sequence becomes available, the true extent to which these unigenes of *S. polycystum* in this study either represent novel proteins, or whether of them are derived from untranslated regions, can be revealed. This observation may reflect less exploration on biological functions on seaweeds and the scarcity of relevant information within the protein database.

Among the *S. polycystum* ESTs, the major molecular functions were binding and catalytic activities, which are also found in other algae (Liu *et al*., 2014; Wang *et al*., 2014a). However, genes expressed under special conditions such as in the protoplast stage are probably missing.

The most abundant unigenes/protein from *S. polycystum* are associated with P-loop containing nucleoside triphosphate hydrolase superfamily according to databases SUPERFAMILY and CATH-Gene3D. The P-loop nucleoside triphosphate hydrolase (NTPase) is the most prevalent domain of several distinct nucleotide-binding protein folds. Members of the P-loop NTPase domain superfamily are defined by a conserved

nucleotide phosphate-binding motif, which is also referred as Walker A and Walker B motif that bind the beta-gamma phosphate moiety of the bound nucleotides and Mg^{2+} cation (Macovei *et al*., 2012).

NAD(P)-binding Rossman-fold domains are the second abundant proteins in *S. polycystum* according to databases IPR and SUPERFAMILY. NAD(P)-binding domain belongs to a large protein family and many different enzymes contain it. It is an important domain in every organism. Anderson *et al*. (2015) exploited algal NADPH oxidase for biophotovoltaic energy, and showed that plasma membrane NADPH oxidase activity is a significant component of light-dependent generation of electricity by the unicellular green alga *Chlamydomonas reinhardtii*. Glyceraldehyde-3-phosphate dehydrogenase is regulated by ferredoxin-NADP reductase in the diatom *Asterionella formosa* (Mekhalfi *et al*., 2014).

It is worth to note that thioredoxin-like fold domain proteins are highly abundant Interpro domain in *S. polycysum*. Thioredoxin (Trx) is an important redox active protein (12 kDa) that regulates the cellular reduction/oxidation status and various important cellular functions, such as reducing protein disulfides, protecting cells from oxidative stress, cell proliferation, signal transduction, transcription regulation, enzymatic detoxification of xenobiotics and it also maintains reductive homeostasis in the cytosol by reducing incorrect disulfides that have formed in proteins, thus returning these proteins to their native state (Karlenius & Tonissen, 2010; Chi *et al*., 2013). The over-representation of these stress-associated domains in *S. polycystum* ESTs indicates that these domains are essential for the survival for *S. polycystum*, which always in the stress conditions (Hanelt & Nultsch, 2012).

5.4 Interesting Transcripts Identified in *S. polycystum*

Alginate is the main structural polysaccharides in brown algae which cementing the cells together. It is located in the matrix of the algae and also in the cell wall (Bourgougnon & Stiger-Pouvreau, 2012). In the pathway shown in Figure 2.2, the conversion GDP-mannose to GDP-mannuronic acid, GDP-mannuronic acid to mannuronan, and to the final C5-epimerization of the polymannuronan chain into alginic acid are specific to alginate biosynthesis (Nyvall *et al*., 2003). In this study, one transcript encoding phosphomannomutase (PMM) (Contig57), two transcripts encoding GDP-mannose 6-dehydrogenase (GMD) (SP01705 & SP00900, respectively) and three transcripts (SP01411 (MC5E-1), SP01750, and SP02271 (MC5E-2)) were identified from the *S. polycystum* ESTs. However, the transcripts encoding mannuronan synthase (MS) were not identified among the ESTs for *S. polycystum* in this study. This could probably due to the fact that cell-wall polysaccharide biosynthesis may not have been actively taking place in the starting materials, or the number of ESTs was too few for the identification of ESTs encoding MS. One transcript encode MC5E (SP01750) has frameshift possibly due to sequencing error or mutation, resulting in a completely different translation from the original, thus only two transcripts were used for the subsequent analysis.

A large multigenic family of enzymes homologous to bacterial MC5Es catalyze the last step of alginate biosynthesis were identified in *L. digitata* (Nyvall *et al*., 2003), *E. siliculosus* (Cock *et al*., 2010; Michel *et al*., 2010b), *Discosporangium mesarthrocarpum*, *Schizocladia ischiensis* (Yamagishi *et al*., 2014) and brown algae from the OneKP which including *Colpomenia sinuosa*, *Desmarestia viridis*, *Dictyopteris undulata*, *S. fusiforme*, *Ishige okamurai*, *Saccharina sculpera*, *Petalonia fascia*, *Saccharina japonica*, *S. hemiphyllum* var. *chinense*, *S. henslowianum*, *S.*

integerrimum, *S. thunbergii*, *S. vachellianum*, *S. horneri*, *S. muticum*, *Scytosiphon lomentaria* and *Scytosiphon dotyi* (Wang *et al*., 2014b). Nyvall *et al*. (2003) and Wang *et al*. (2014b) proposed that brown algae and bacterial MC5E, or maybe the entire alginate biosynthesis pathway of brown algae, have arisen from common ancestor or maybe from an ancestral alginate-producing bacterium by horizontal gene transfer based on their structural family and phylogenetics analyses. It is not sure whether the MC5Es or the entire alginate biosynthesis pathway of *S. polycystum* may also is from an ancestral alginate-producing bacterium. This however, has to be confirmed by the detailed studies on the relationship between MC5Es of *S. polycystum* and bacterial MC5Es.

ClustalW alignment showed that these two MC5E encoding sequences, SP01411 (MC5E-1) and SP02271 (MC5E-2), from *S. polycystum* contain 48 % of peptide similarity (identity). However, it is not sure if some of these conserved sequences are important for the enzyme function. In this study, these two partial transcripts encoding MC5E ezymes sequences were very short, and were not functionally classified by GO, EC and thus KEGG analysis.

It is of interest that the putative transcripts encoding for enzymes for fructose and mannose metabolism are shown in Table 4.7. Two transcripts encoding GDP-D-mannose 4,6-dehydratase (GM46D) (Contig159, Contig254), and a transcript encoding GDP-4-keto-6-deoxy-d-mannose-3,5-epimerase-4-reductase (GFS) (Contig87), which are involved in sulfated fucose biosynthesis in *S. polycystum* were identified. However, the putative transcripts encoding for fucosyltransferases and sulfotransferases were not identified in *S. polycystum*. This could likely due to their low abundance and/or the gene sequences are not similar enough to its homologoues in the public databases. One of the transcripts encoding GM46D (Contig254) has frameshift maybe due to sequencing error

or mutation, thus only one trancript encode GM46D (Contig159) was used for the subsequent analysis. Yamagishi *et al*. (2014) have reported that *D. mesarthrocarpum* has an entire pathway for sulfated fucam metabolism, which is similar to that of *E. siliculosus*. They also reported that some genes related to the metabolism of sulfated fucans were absent from their transcriptome data sets of *D. mesarthrocarpum*, *Schizocladia ischiensis* and *Phaeothamnion confervicola*, respectively, and it was not possible to elucidate whether this was due to the low expression of the genes, or absence from the genome. These transcripts are vital for the understanding of fucoidan biosynthesis in brown algae, where fucoidan is another main component of brown algae cell wall and so far, the fucoidan metabolism has yet to be characterised (Michel *et al*., 2010b). There are two pathways for GDP-L-fucose biosyntheis in brown algae, namely *de novo* pathway and the salvage pathway, were found in *E. siliculosus* (Michel *et al*., 2010b) and in *S. thunbergii* (Liu *et al*., 2014). In this study, a transcript encoding GM46D and a transcript encoding GFS have been identified, indicating that GDP-fucose of *S. polycystum* can produced via the *de novo* pathway. Wong *et al*. (2007) also had identified partial cDNAs encoding GM46D and fucokinase. We did not find any evidence of subcellular localization for GM46D-Contig159 and GFS-Contig87 in specific cellular compartments (no known N-terminal target peptides are detected using HECTAR). This is in agreement with the presence of these proteins in cytoplasms, e.g. as reported in *Plasmodium falciparum* (Sanz *et al*., 2013).

Brown algae and other Stramenopiles exhibit some remarkable metabolic characteristic, notably in their unique carbon (C) storage metabolism (Gravot *et al*., 2010). Indeed, brown algae do not use the photoassimilate D-fructose-6-phosphate (F6D) to produce starch, saccharose and fructans as in most of the higher plants, but instead produce the β-1,3-glucan laminarin and mannitol (Bonin *et al*., 2015; Michel *et*

al., 2010a). In this study, three transcripts encoding M1PDH (Contig109, SP01050 and SP01780) were identified in *S. polycystum*, however, the putative genes encoding for M1Pase, M2DH and HK were not identified in *S. polycystum*. The putative genes encoding M1PDH, M2DH and HK were identified in *S. thunbergii* transcriptome (Liu *et al*., 2014). Liu *et al*. (2014) concluded that *S. thunbergii* probably has a similar mannitol metabolism process as *E. siliculosus*. Future efforts should be taken to elucidate the mannitol metabolism process in *S. polycystum*.

Transcripts encoding for enzymes involved in fructose and mannose metabolism which also identified in this study, including fructose-bisphosphate aldolase (FBA), ribose-5-phosphate isomerase (RPI), UDP-glucose 6-dehydrogenase, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and triosephosphate isomerase (TPI). In *Phaeodactylum tricornutum*, class I plastid FBA and chromalveolate-specific class II plastid FBA are localized in the pyrenoid region of the chloroplast, playing a significant role in regulation of the diatom carbon concentrating mechanism. Where, cytosolically localized FBAs in *P. tricornutum* were likely play a role in glycolysis and cytoskeleton function (Allen *et al*., 2012).

Marine organisms living in the intertidal or shallow subtidal habitats are regularly exposing to recurring, harsh changes biotic and physical abiotic factors in the environment (Thomsen & Melzner, 2010). Seaweeds are usually immobile, thus they have to rely on cellular mechanisms to tolerate stress in order to survive in the intertidal zone (Dittami *et al*., 2009). In this study, there are a number of stress-related, defense and/ or detoxification transcripts were identified: ascorbate peroxidase (APX), cytochrome c peroxidase, glutamine synthetase, glutathione S-transferase (GST), glutathione reductase (GR), glutathione peroxidase (GPX) and superoxide dismutase (SOD) (Table 4.7). APX has been reported to be involved in the scavenging in many

organisms (Dabrowska *et al*., 2007; Caverzan *et al*., 2012; Urzica *et al*., 2012). SOD also is widely assumed to play a role in the detoxification of ROS which caused by enviromental stresses (Shiraya *et al*., 2015). Bowler *et al*. (1992) suggested that glutathione reductase may cooperate with SOD to remove superoxide radicals. Kumar *et al*. (2010) suggested that the antioxidant enzymes (APX, GR and SOD), phycobiliproteins and mineral nutrients have potential roles to combat the oxidative stress which induced by the hypo- and hyper-salinity treatments study in *Gracilaria corticata*. Another group of stress proteins were heat shock proteins (HSPs), with a total of 47 EST sequences (Table 4.7) were identified among the *S. polycystum* ESTs. Members were found of all the five major classes of plant HSPs, including HSP 20, HSP 40, HSP 60, HSP 70 and HSP 90. HSPs protect cells against many stresses The heat shock proteins aid in the folding of damaged as well as in the newly formed proteins (Timperio *et al*., 2008; Horváth *et al*., 2012). Plants synthesize both high molecular mass HSPs (from 60 to 110 kDa) and small HSPs (from 15 to 45 kDa) when they exposed to high temperatures stress (Timperio *et al*., 2008).

There has not been any experimental comparison of candidate reference genes to obtain the most suitable for *S. polycystum* to date. This present work is the first study on the expression stability of a set of reference genes that aims to identify a set of control genes for normalization of transcript levels in *Sargassum*. In this study, a few candidate genes among the *S. polycystum* ESTs were selected. GeNorm pairwise analysis was performed to test the robustness of the data in this study and AT, RPL3A and EF1A were among the most stable genes in the sample sets under study. The gene encoding RPS15A had a greater variation in gene expression or the least stable in this study. The gene encoding EF1A was also reported to be one the best reference genes for sugarcane, *Saccharum* spp. (Guo *et al*., 2014) and fungus, *Trichoderma reesei* (Dashtban *et al*.,

2012). The gene encoding RPL3A was reported to be one of the best reference gene for *Arabidopsis thaliana* (Vellosillo *et al*., 2013). The AT has been used in studies where AT has been identified as variable gene (Zhu *et al*., 2013), however in this study, AT is one of the most stable genes.

Little is known at the molecular level in brown algae, while alginate biosynthesis has been well investigated in bacteria with identification of genes encoding mannuronan C5-epimerase, which involved in the alginate biosynthesis and in their regulation (Castillo *et al*., 2013; Flores *et al*., 2013). In the *Ectocarpus* genome, genes encoding mannuronan C5-epimerase were highly represented, but none of these genes were among the most significantly regulated loci, although several of them were either induced or repressed under different stress conditions (Dittami *et al*., 2009). In this study, it is important to identify which transcripts encoding MC5E are expressed during rainy or dry season in order to understand the genetic control of alginate biosynthesis. This study showed that the transcript level of MC5E-1 (SP01411) and MC5E-2 (SP02271) were up- and down-regulated, respectively during rainy season (July 2010) in relative to dry season (December 2010). The identity of these two partial transcript sequences are only 48 %. Thus, they might encode for the enzymes that involved in generation of different alginates or play different roles during alginate biosynthesis. Tonon *et al*. (2008) supported the hypothesis that under pathogen attack, specific MC5E genes are formed, based on the accumulation of MC5E transcripts during sporophyte elicitation and protoplast regeneration, where in response to environmental stimuli, *L. digitata* modified its cell wall. The expression of genes encoding mannuronan C5-epimerase, according to Nyvall *et al*. (2003), was up-regulated in spring and winter. Roeder *et al*. (2005) found that genes encoding mannuronan C5-epimerase are highly expressed in *L. digitata* protoplasts. Three transcripts encoding MC5Es in this study

may reflect the important of alginate biosynthesis in *S. polycystum.* However, more validation is needed before definitive conclusions can be made. In this study, alginate extraction from *S. polycystum* was not conducted due to insufficient samples.

In this study, the relative abundance of one GM46D and one GFS genes at rainy versus dry seasons were studied. However, the transcript abundance of Contig159 (GM46D) declined 0.3-fold during rainy season, while the transcript abundance of Contig87 (GFS) was increased 0.2-fold during rainy season compared to dry season. Both genes were exhibited below 2-fold change, which no significant difference between two seasons for both genes. Thus, more genes should be included in the future RT-qPCR or microarray study in order to get a more significant result. Wang *et al*. (2014a) used the RPKM to quantitate the global gene expression level of six *Sargassum* species and it showed that *S. fusiforme* among the six *Sargassum* showed more genes expressed for genes encoding fructokinase and GM46D than others. They predicted that *S. fusiforme* might accumulate more sulfated polysaccharides than others in such conditions when algal samples were collected. However, they still need further molecular and physiological experiments to confirm this phenomenon. All these information will definitely assist future functional genomic analysis of *S. polycystum* towards a better understanding of their gene functions.

CHAPTER 6: CONCLUSION

6.1 Conclusion

A total of five different RNA extraction methods were examined for RNA extraction from *S. polycystum*. The newly optimized method developed in this study was able to extract high quality RNA from *S. polycystum*. This method successfully increased the RNA yield and eliminated the polysaccharide contamination. Sufficient amount of mRNA was successfully obtained from *S. polycystum* to construct a cDNA library. A primary cDNA library with a titer 1.0 x 10^6 pfu/ml and an amplified cDNA library with a titer of 4.2 x 10^{10} pfu/ml were obtained. A total of 2,577 high-quality readable EST sequences were generated from *S. polycystum*. A total of 1,667 unigenes consisting of 295 contigs and 1429 singletons were obtained. The redundancy of ESTs in *S. polycystum* cDNA library was approximately 45 %. Functional annotation showed interesting genes were found among the ESTs that are involved in alginate and sulfated fucan/ fucose biosynthesis pathways such as mannuronan C5-epimerase (MC5E), GDP-D-mannose 4,6 dehydratase (GM46D) and GDP-4-keto-6-D-mannose-epimerase-4-reductase (GFS). Quantitative real-time PCR (RT-qPCR) analysis revealed differential expression of these transcripts in seaweeds collected during rainy and dry seasons. This library provides a robust sequence resource which will be useful for cloning of full-length transcripts.

6.2 Appraisal of Study

EST approach which is one of the effective and cheaper solutions for complements existing genome projects or for use in gene discovery, across a variety of organisms. Besides, ESTs provide a robust sequencing resource for phylogenetics (Shen *et al*.,

2013), improving genome annotation (Parkinson & Blaxter, 2009), comparative genomics (Soanes & Talbot, 2006), transcript profiling (Kanth *et al*., 2015) and proteomics (Victor *et al*., 2013; Yang *et al*., 2010). At the time of writing (January 2016), numerous organisms have generated large amount of ESTs which including Ochrophyta (85,460 ESTs), Rhodophyta (455,144), Chlorophyta (569,265), Apicomplexa (673,544) and Animalia (43,782,421) were deposited in NCBI database (Guiry & Guiry, 2016).

EST approach has two major advantages compared to genomics. Firstly, as poly(A) tails of bacterial transcripts are too short to be primed by standard oligo-dT primers, thus bacterial contamination of eukaryotic cultures can be avoided. Secondly, starting material in the current study, a brown seaweed, has rigid cell walls which resist digestion with cell wall lysing enzymes, thus harsh mechanical methods for cell wall breakage is required. Smaller mRNAs will still remain intact, while genomic DNA may become too fragmented to go through such treatment for genome sequencing. EST approach also has some advantages over the next generation sequencing technologies. The EST method generates a library which contains the DNA information that can be used for subsequent analysis or confirmation. For example, EST sequences can be aligned to the genome if the genome of the organism has been sequenced. Secondly, the transcripts assembly is easier and do not require expensive and novel alignment algorithms when compared to NGS transcripts assembly. On the other hand, EST have become a tool to refine the predicted transcripts for certain genes, which leads to the prediction of their protein products and ultimately of their function (gene model). However, a major drawback of ESTs technology is that it cannot generate huge amount of data rapidly at a modest cost, which may be offered by the Next-generation sequencing (NGS) approach.

The isolation of high-quality RNA from brown seaweeds has always been problematic due to a high content of polyphenolic compounds and polysaccharides in their tissues compared to higher plants and other organisms. Seaweed tissues have high levels of polysaccharides and secondary metabolites that are released after disruption of cells embedded in viscous polysaccharides. The task is even more difficult due to the enzymatic degradation of RNA molecules by endogenous RNase. Furthermore, the isolation of intact and high-quality RNA without DNA contamination is always important and necessary for gene expression studies including cDNA library construction and RT-qPCR. Besides, numerous methods for the isolation of total RNA from recalcitrant plants and algae are time-consuming, expensive and hence, limiting the processing of a large number of samples. In the present study, we developed a newly optimized RNA extraction method. This new method eliminated most of the interfering molecules efficiently and yielded translucent and water-soluble high quality RNA pellets. It is a rapid, relatively nontoxic, and inexpensive RNA extracton method for *S. polycystum*. The cDNA library in this study prepared from the RNA isolated by this new protocol fulfills two key characteristics of a high quality cDNA library, such as a primary cDNA library containing 1×10^6 recombinant clones was obtained, and secondly, the average size of the cDNA inserts is ~1 kb or greater, and thus giving a high probability of isolating the full-length cDNAs (Sambrook & Russell, 2001).

6.3 Areas for Future Research

1) Isolate the full-length cDNAs of mannuronan C5-epimerase (MC5E) from *S. polycystum*, and followed by characterization of these MC5E genes.

2) Whole genome sequencing using PacBio or Illumina technologies, of additional species such as *S. polycystum*, *S. binderi* and *S. siliquosum* to compare their genomic variation, structural variants, haplotypes and epigenetics, should be conducted.
3) Gene expression of *Sargassum* spp. by using RNA sequencing (RNA-seq) technology should be used to reveal or determine how the transcriptomes are affected under different conditions (such as different salinity (ppt), underwater irradiance (photon flux density, µmol photon m^{-2} s^{-1}), seawater temperature (°C) or different life stages (protoplast, juvenile, matured).
4) Transcripts related to the alginate and sulphated fucan biosynthesis pathways should be further studied, to enable us to increase the content of the metabolites by genetic engineering in the future.

REFERENCES

Abbas, S. H., Ismail, I. M., Mostafa, T. M. & Sulaymon, A. H. (2014). Biosorption of heavy metals: a review. *Journal of Chemical Science and Technology, 3*, 74-102.

Accerbi, M., Schmidt, S. A., De Paoli, E., Park, S., Jeong, D. H., & Green, P. J. (2010). Methods for isolation of total RNA to recover miRNAs and other small RNAs from diverse species. *Methods in Molecular Biology, 592*, 31-50.

Adams, M.D., Kelley, J.M., & Gocayne, J.D. (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. *Science, 252*, 1651-1656.

Adams, M.D., Celniker, S.E., Holt, R. A., …Venter, J. C. (2000). The genome sequence of *Drosophila melanogaster, Science*, *287*, 2185-2195.

Addgene. (2016). Antibiotic concentrations for bacterial selection. Retrieved on 13 February, 2016 at https://www.addgene.org/mol-bio-reference/antibiotics/

Ahmed, E. M. (2015). Hydrogel: preparation, characterization, and applications: a review. *Journal of Advances Research, 6*, 105-121.

Ajisaka, T., Phang, S. M., & Yoshida, T. (1999). Preliminary report of *Sargassum* species collected from Malaysian coast. In I.A Abbott (Ed.), *Pacific Species VII* (pp. 23-42). California, La Jolla: California Sea Grant College, University of California.

Ale, M. T., & Meyer, A. S. (2013). Fucoidans from brown seaweeds: an update on structures, extraction techniques and use of enzymes as tools for structural elucidation. *RSC Advances, 3*, 8131-8141.

Ale, M. T., Mikkelsen, J. D., & Meyer, A. S. (2011). Important determinants for fucoidan bioactivity: a critical review of structure-function relations and extraction methods for fucose-containing sulfated polysaccharides from brown seaweeds. *Marine Drugs, 9*, 2106-2130.

Alibaba.com. (2016). Retrived on 10 January 2016 from http://www.alibaba.com/showroom/sodium-alginate-price.html

Allen, A. E., Moustafa, A., Montsant, A., Eckert, A., Kroth, P. G., & Bowler, C. (2012). Evolution and functional diversification of fructose bisphosphate aldolase genes in photosynthetic marine diatoms. *Molecular Biology and Evolution, 29*, 367-379.

Al-Swailem, A. M., Shehata, M. M., Abu-Duhier, F. M., Al-Yamani, E. J., Al-Busadah, K. A., Al-Arawi, M. S., … Out, H. H. (2010). Sequencing, analysis, and annotation of expressed sequence tags for *Camelus dromedarius*. *PLoS ONE, 5*, e10720.

Andersen, C. L., Ledet-Jensen, J., & Ørntoft, T. (2004). Normalization of real-time quantitative RT-PCR data: a model based variance estimation approach to identify genes suited for normalization – applied to bladder and colon – cancer data – sets. *Cancer Research, 64*, 5245-5250.

Anderson, A., Laohavisit, A., Blaby, I. K., Bombelli, P., Howe, C. J., Merchant, S. S., … Smith, A. G. (2016). Exploiting algal NADPH oxidase for biophotovoltaic energy. *Plant Biotechnology Journal, 14*, 22-28.

Ang, P. (2007). Phenology of *Sargassum* spp. in Tung Ping Chau marine Park, Hong Kong SAR, China. In R. Anderson, J. Brodie, E. Onsøyen, & A. T., Critchley. (Eds.), *Eighteenth International Seaweed Symposium:Proceeding of the 18th International Seaweed Symposium in Bergen, Norway* (pp. 403-410). Netherlands: Springer International Publishing AG.

Apostolidis, E., & Lee, C. M. (2012). Brown seaweed-derived phenolic phytochemicals and their biological activities for functional food ingredients with focus on *Ascophyllum nodosum*. In S.K. Kim (Ed.), *Handbook of marine macroalgae: Biotechnology and applied phycology* (pp. 356-370). United Kingdom, UK: John Wiley & Sons, Ltd.

Apt, K. E., Clendennen, S. K., Powers, D. A. & Grossman, A. R. (1995). The gene family encoding the fucoxanthin chlophyll proteins from the brown alga *Macrocystis pyrifera*. *Molecular & General Genetics*, *246*, 455-464.

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

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*, 333-338.

Bartsch, I., Wiencke, C., Bischof, K., Buchholz, C.M., Buck, B.H., Eggert, A., Feuerpfeil, P., … Wiese, J. (2008). The genus *Laminaria* sensu lato: recent insights and developments. *European Journal of Phycology, 43*, 1-86.

Battacharyya, D., Babgohari, M. Z., Rathor, P. & Prithiviraj, B. (2015). Review: Seaweed extracts as biostimulants in horticulture. *Scientia Horticulture, 196*, 39-48.

Berteau, O., & Mulloy, B. (2003). Sulfated fucans, fresh perspectives: structures, functions, and biological properties of sulfated fucans and an overview of enzymes active toward this class of polysaccharide. *Glycobiology, 13*, 29R-40R.

Bhattacharya, D., Price, D. C., Chan, C. X., Qiu, H., Rose, N., Ball, S., … Yoon, H. S. (2013). Genome of the red alga *Porphyridium purpureum*. *Nature Communications, 4*, 1941.

Bilan, M. I., Grachev, A. A., Shashkov, A. S., Thuy, T. T. T., Van, T. T. T., Ly, B. M., … Usov, A. I. (2013). Preliminary investigation of a highly sulfated galactofucan fraction isolated from the brown alga *Sargassum polycystum*. *Carbohydrate Research, 377*, 48-57.

Bisht, K. K., Dudognon, C., Chang, W. G., Sokol, E. S., Ramirez, A., & Smith, S. (2012). GDP-mannose-4,6-dehydratase is a cytosolic partner of tankyrase 1 that inhibits its poly(ADP-ribose) polymerase activity. *Molecular and Cell Biology, 32*, 3044-3053.

Bixler, H. J. & Porse, H. (2011). A decade of change in the seaweed hydrocolloids industry. *Journal of Applied Phycology, 23*, 321-335.

Boateng, J. S., Matthews, K. H., Stevens, H. N. E., & Eccleston, G. M. (2008). Wound healing dressings and drug delivery systems: a review. *Journal of Pharmaceutical Sciences, 97*, 2892-2923.

Bonin, P., Groisillier, A., Raimbault, A., Guibert, A., Boyen, C., & Tonon, T. (2015). Molecular and biochemical characterization of mannitol-1-phosphate dehydrogenase from the model brown alga *Ectocarpus* sp. *Phytochemistry, 117*, 509-520.

Bourgougnon, N., & Stiger-Pouvreau, V. (2012). Chemodiversity and bioactivity within red and brown macroalgae along the French coasts, metropole and overseas departments and territories. In S.K. Kim (Ed.), *Handbook of marine macroalgae: Biotechnology and applied phycology* (pp. 500-521). United Kingdom, UK: John Wiley & Sons, Ltd.

Bowler, C., Allen, A. E, Badger, J. H., Grimwood, J., Jabbari, K., Kuo, A., … Grogoriev, I. V. (2008). The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature, 456*, 239–244.

Bowler, C., Montagu, M. V., & Inzé, D. (1992). Superoxide dismutase and stress tolerance. *Annual Review of Plant Pgysiology and Plant Molecular Biology, 43*, 83-116.

Brownlee, I. A., Seal, C. J., Wilcox, M., Dettmar, P. W., & Pearson, J. P. (2009). Applications of alginates in food. In B.H.A. Rehm (Ed.), *Alginates: Biology and Applications, Microbiology Monographs 13* (pp. 211-228). Berlin, Heidelberg: Springer.

Bugos, R.C., Chiang, V. L., Zhang, X. H., Campell, E. R., Podilla, G. K., & Campell, W. H. (1995). RNA isolation from plant tissues recalcitrant to extraction in guanidine. *Biotechniques, 19*, 734-737.

Bustin, S. A. (2010). Why the need for qPCR publication guidelines? – The case for MIQE. *Methods, 50*, 217-226.

Caffall K. H. & Mohnen D. (2009). The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydrate Research*, *344*, 1879-1900.

Cao, R. A., Lee, Y., & You, S. G. (2014). Water soluble sulfated-fucans with immune-enhancing properties from *Ecklonia cava*. *International Journal of Biological Macromolecules, 67*, 303-311.

Caverzan, A., Passaia, G., Rosa, S. B., Ribeiro, C. W., Lazzarotto, F., & Margis-Pinheiro, M. (2012). Plant responses to stresses: Role of ascorbate peroxidase in the antioxidant protection. *Genetic and Molecular Biology, 35*, 1011-1019.

Castillo, T., Heinzle, E., Peifer, S., Schneider, K., & Peña, M. C. F. (2013). Oxygen supply strongly influences metabolic fluxes, the production of poly(3-hydroxybutyrate) and alginate, and the degree of acetylation of alginate in *Azotobacter vinelandii*. *Process Biochemistry, 48*, 995-1003.

Chan, C. X., Bhattacharya, D., & Reyes-Prieto, A. (2012a). Endosymbiotic and horizontal gene transfer in microbial eukaryotes. *Mobile Genetic Elements, 2*, 101-105.

Chan, C. X., Blouin, N. A., Zhuang, Y., Zäuner, S., Prochnik, S. E., Lindquist, E., … Bhattacharya, D. (2012b). *Porphyra* (Bangiophyceae) transcriptomes provide insights into red algal development and metabolism. *Journal of Phycology*, *48*, 1328–1342.

Chan, C. X., Reyes-Prieto, A., & Bhattacharya, D. (2011). Red and green algal origin of diatom membrane transporters: Insights into environmental adaptation and cell evolution. *PLoS ONE, 6*, e29138.

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

Chan, K. L., Ho, C. L., Parameswari Namasivayam & Suhaimi Napis. (2007). A simple and rapid method for RNA isolation from plant tissues with high phenolic compounds and polysaccharides. *Protocol Exchange*. doi:10.1038/nprot.2007.184.

Chi, Y. H., Paeng, S. K., Kim, M. J., Hwang, G. Y., Melencion, S. M. B., Oh, H. T., & Lee, S. Y. (2013). Redox-dependent functional switching of plant proteins accopanying with their structural changes. *Frontiers in Plant Science, 4*, 277.

Christaki, E., & Bonos, E., Giannenas, I., Florou-Paneri, P. (2013). Functional properties of carotenoids originating from algae. *Journal of the Science of Food and Agriculture, 93*, 5-11.

Chomczynski, P., & Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry, 162*, 156-159.

Cinar, M. U., Islam, M. A., Pröll, M., Kocamis, H., Tholen, E., Tesfaye, D., … Uddin, M. J. (2013). Evaluation of suitable reference genes for gene expression studies in porcine PBMCs in response to LPS and LTA. *BMC Research Notes, 6*, 56.

Cock, J.M., Sterck, L., Rouzé, P., Scornet, D., Allen, A. E., Amoutzias, G., … Winker, P. (2010). The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. *Nature, 465*, 617-621.

Coelho, S. M., Scornet, D., Rousvoal, S., Peters, N. T., Dartevelle, L., Peters, A. F., & Cock, J. M. (2012). *Ectocarpus*: a model organism for the brown algae. *Cold Spring Harbor Protocols, 2012*.

Cofrades, S., López-López, I., & Jiménez-Colmenero, F. (2012). Applications of seaweed in meat-based functional foods. In S.K. Kim (Edi.), *Handbook of Marine Macroalgae: Biotechnology and applied phycology* (pp. 491-499). United Kingdom, UK: John Wiley & Sons, Ltd.

Collén, J., Guisle-Marsollier, I., Leger, J. J., & Boyen, C. (2007). Response of the transcriptome of the intertidal red seaeed *Chondrus crispus* to controlled and natural stresses. *New Phytologist, 176*, 45-55.

Collén, J., Porcel, B., Carré, W., Ball, S. G., Chaparro, C., Tonon, T., … Boyen, C. (2013). Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. *Proceedings of the National Academy of Sciences, 110*, 5247-5252.

Collombo-Pallotta, M. F., García-Mendoza, E., & Ladah, L. B. (2006). Photosynthetic performance, light absorption, and pigment composition of *Macrocystis pyrifera* (Laminariales, Phaeophyceae) blades from different depths. *Journal of Phycology, 42*, 1225-1234.

Conesa, A, & Götz, S. (2008). Blast2GO: A comprehensive suite for functional analysis in plant genomics. *International Journal of Plant Genomics*, 2008, 619832.

Costa, V., Angelini, C., De Feis, I., & Ciccodicola, A. (2010). Uncovering the complexity of transcriptomes with RNA-seq. Review article. *Journal of Biomedicine and Biotechnology*. Vol 2010, article ID 853916, 19 pages.

Crépineau, F., Roscoe, T., Kaas, R., Kloareg, B., & Boyen, C. (2000). Characterisation of complementay DNAs from the expressed sequence tag analysis of life cycle stages of *Laminaria digitata* (Phaeophyceae). *Plant Molecular Biology, 43*, 503-513.

Dabrowska, G., Kata, A., Goc, A., Szechyńska-Hebda, M., & Skrzypek, E. (2007). Characteristics of the plant ascorbate peroxidase family. *Acta Biologica Cracoviensia Series Botanica, 49*, 7-17.

Dashtban, M., & Qin, W. (2012). Overexpression of an exotic thermotolerant β-glucosidase in *Trichoderma reesei* and its significant increase in cellulolytic activity and saccharification of barley straw. *Microbial Cell Factories, 11*, 63.

de Castro, E., Sigrist, C. J. A., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P. S., Gasteiger, E., Bairoch, A., & Hulo, N. (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Research*, *34*, W362-W365.

Delmar, K., & Bianco-Peled, H. (2016). Composite chitosan hydrogels for extended release of hydrophobic drugs. *Carbohydrate Polymers, 136*, 570-580.

de Oliveira, L. S., Gregoracci, G. B., Silva, G. G. Z., Salgado, L. T., Filho, G. A., Alves-Ferreira, M., ...Thompson, F. L. (2012). Trancriptomic analysis of the red seaweed *Laurencia dendroidea* (Florideophyceae, Rhodophyta) and its microbiome. *BMC Genomics, 13*, 487.

Derveaux, S., Vandesompele, J., & Hellemans, J. (2010). How to do successful gene expression analysis using real-time PCR. *Methods, 50*, 227-230.

de Széchy, M. T. M., Guedes, P. M., Baeta-Neves, M. H., & Oliveira, E. N. (2012). Verification of *Sargassum natans* (Linnaeus) Gaillon (Heterokontophyta: Phaeophyceae) from the Sargasso Sea off the coast of Brazil, western Atlantic Ocean. *Check List, 8*, 638-641.

Dettmar, P. W., Strugala, V., & Richardson, J. C. (2011). The key role of alginates play in health. *Food Hydrocolloids, 25*, 263-266.

Dillehay, T. D., Ramirez, C., Pino, M., Collins, . B., Rossen, J., & Pino-Navarro, J. D. (2008). Monte Verde: seaweed, food, medicine, and the peopling of South America. *Science, 320*, 784-786.

Dittami, S. M. (2010). Abiotic stress response in *Ectocarpus siliculosus*: a global approach. (Unpublished doctoral dissertation). Université De Rennes 1.

Dittami, S. M., Scornet, D., Petit, J-L., Ségurens, B., Da Silva, C., Corre, E., … Tonon, T. (2009). Global expression analysis of the brown alga *Ectocarpus siliculosus* (Phaeophyceae) reveals large-scale reprogramming of the transcriptome in response to abiotic stress. *Genome Biology, 10*: R66.

Dore, C. M. P. G., Alves, M. G. d. C. F., Will, L. S. E. P., Costa, T. G., Sabry, D. A., Rêgo, L. A. R. d. S., ... Leite, E. L. (2013). A sulfated polysaccharide, fucans, isolated from brown algae *Sargassum vulgare* with anticoagulant, antithrombotic, antioxidant and anti-inflammatory effects. *Carbohydrate Polymers, 91*, 476-475.

Draget, K. I., Smidsrød, O., & Skjåk-Braek, G. (2005). Alginates from algae. In A. Steinbüchel & S.K. Rhee (Eds.), *Polysaccharides and Polyamides in the Food Industry, Properties, producton, and patents* (pp. 1-30). Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.

Ehrig, K., & Alban, S. (2015). Sulfated galactofucan from the brown alga *Saccharina latissima* – variability of yield, structural composition and biosctivity. *Marine Drugs, 13*, 76-101.

Elnashar, M. M., Yassin, M. A., Moneim, A. E. A., & Bary, E. M. A. (2010). Surprising performance of alginate beads for the release of low-molecular-weight drugs. *Journal of Applied Polymer Science, 116*, 2487-3110.

Erro, E., Bundy, J., Massie, I., Chalmers, S., Gautier, A., Gerontas, S., ... Selden, C. (2013). Bioengineering the liver: scale-up and cool chain delivery of the liver cell biomass for clinical targeting in a bioartificial liver support system. *Bioresearch Open Access, 2*, 1-11.

Farbo, M. G., Urgeghe, P. P., Fiori, S., Marceddu, S., Jaoua, S., & Migheli, Q. (2016). Adsorption of ochratoxin A from grape juice by yeast cells immobilised in calcium alginate beads. *International Journal of Food Microbiology, 217*, 29-34.

Fasihuddin B. A. & Siti H. S. (1994). Characterization of alginate from various brown seaweeds. *Algae Biotechnology in the Asia-Pacific Region* (pp. 156-163). Kuala Lumpur, Malaysia: University of Malaya.

Firdaus, M., Nurdiani, R., & Prihanto, A. A. (2015). Antihyperglycemic of *Sargassum* sp extract. In S.K. Kim & K. Chojnacka (Eds.), *Marine algae extracts: Processes, products, and applications, 2 volume set* (pp. 381-394). United Kingdom, UK: John Wiley & Sons, Ltd.

Flores, C., Moreno, S., Espín, G., Peña, C., & Galindo, E. (2013). Expression of alginases and alginate polymerase genes in response to oxygen, and their relationship with the alginate molecular weight in *Azotobacter vinelandii*. *Enzyme and Microbial Technology, 53*, 85-91.

Flöthe, C. R., Molis, M., & John, U. (2014). Induced resistance to periwinkle grazing in the brown seaweed *Fucus vesiculosus* (Phaeophyceae): molecular insights and seaweed-mediated effects on herbivore interactions. *Journal of Phycology, 50*, 564-576.

Fong, P. & Paul, V. J. (2011). Coral reef algae. In Z. Dubinsky & N. Stambler (Eds.), *Coral reefs: an ecosyctem in transition* (pp. 241-272). Netherlands: Springer.

Food and Agriculture Organization of the United Nations (FAO). (2014). The state of world fisheries and aquaculture: opportunities and challenges. (223 p.) Rome. Retrieved from http://www.fao.org/3/a-i3720e.pdf

Food and Agriculture Organization of the United Nations (FAO). (2016). Fishery Statistical Collections: Global Production Statistics 1950-2013. Retrieved on 15 January 2016 from http://www.fao.org/fishery/statistics/en

Gao, W. R., Wang, X. S., Liu, Q. Y., Peng, H., Chen, C., Li, J. G., … Ma, H. (2008). Comparative analysis of ESTs in response to drought stress in chickpea (*C. arietinum* L.). *Biochemical and Biophysical Research Communications, 376*, 578-583.

Gehrig, H. H., Winter, K., Cuishman, J., Borland, A., & Taybi, T. (2000). An improved RNA isolation method for succulent plant species rich in polyphenols and polysaccharides. *Plant Molecular Biology Reporter, 18*, 369-376.

Ghosh, R., Banerjee, K., & Mitra, A. (2012). Eco-biochemical studies of common seaweeds in the lower gangetic delta. In S.K. Kim (Ed.), *Handbook of marine macroalgae: Biotechnology and applied phycology* (45-57). United Kingdom, UK: John Wileys & Sons, Ltd.

Gill, S. S., & Tuteja, N. (2010). Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiology and Biochemistry, 48*, 909-930.

Gimeno, J., Eattock, N., Van Deynze, A., & Blumwald, E. (2014). Selection and validation of reference genes for gene expression analysis in switchgrass (*Panicum virgatum*) using quantitative real-time RT-PCR. *PLoS One, 9*, e91474.

Goffeau, A, Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., … Oliver, S. G. (1996). Life with 6000 genes. *Science, 274*, 563-567.

Goh, P. N. (2008). Determination of the nutritional and bioactive properties of selected marine brown algae (Phaeophyta) and the potential as a herbal tea (Master's thesis, University Malaysia Sabah).

Goossens, K., Van Poucke, M., Van Soom, A., Vandesompele, J., Van Zeveren, A., & Peelman, L. J. (2005). Selection of reference genes for quantitative real-time PCR in bovine preimplantation embryos. *BMC Developmental Biology, 5*, 27.

Gravot, A., Dittami, S. M., Rousvoal, S., Lugan, R., Eggert, A., Collén, J., …Tonon, T. (2010). Diurnal oscillations of metabolite abundance and gene analysis provide new insights into central metabolic processes of the brown alga *Ectocarpus siliculosus*. *New Phytologist, 188*, 98-110.

Groisillier A., Shao Z., Michel G., Goulitquer S., Bonin P., Krahulec S., Nidetzky B., Duan D., Boyen C. & Tonon T. (2014). Mannitol metabolism in brown algae involves a new phosphatase family. *Journal of Experimental Botany, 65*, 559-570.

Gschloessl B., Guermeur Y., & Cock J. M. (2008). HECTAR: a method to predict subcellular targeting in heterokonts. BMC Bioinformatics 9: 393.

Guiry, M. D., & Guiry, G. M. (2016). AlgaeBase. World-wide electronic publication, National University of Ireland, Galway**.** Retrieved from http://www.algaebase.org

Guo, J., Ling, H., Wu, Q., Xu, L., & Que, Y. (2014). The choice of reference genes for assessing gene expression in sugarcane under salinity and drought stresses. *Scientific Reports, 4*, 7042.

Gupta, A. K., & Gupta, U. D. (2014). Next generation sequencing and its applications. In A. Verma & A. Singh (Eds.), *Animal Biotechnology* (pp. 345-367). Elsevier: Academic Press.

Hahn, T., Lang, S., Ulber, R., & Muffler, K. (2012). Novel procedures for the extraction of fucoidan from brown algae. *Process Biochemistry, 47*, 1691-1698.

Hall, T. (1999). Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium, 41*, 95-98.

Hanelt, D., & Nultsch, W. (2012). Photoinhibition in seaweeds. In G. Heldaier & D. Werner (Eds.), *Environmental Signal Processing and Adaptation* (pp. 141-168). Berlin, Germany: Springer Science & Business Media

Harley, C. D. G., Anderson, K. M., Demes, K. W., Jorve, J. P., Kordas, R. L., Coyle, T. A., & Graham, M. H. (2012). Effects of climate change on global seaweed communities. *Journal of Phycology, 48*, 1064-1078.

Harrow, J., Nagy, A., Reymond, A., Alioto, T., Patthy, L., Antonarakis, S. E., Guigó, R. (2009). Identifying protein-coding genes in genomic sequences. *Genome Biology, 10*: 201.

Hazra, M., Mandal, D. D., Mandal, T., Bhuniya, S., & Ghosh, M. (2015). Designing polymeric microparticulate drug delivery system for hydrophobic drug quercetin. *Saudi Pharmaceutical Journal, 23*, 429-436.

Heinrich, S., Frickenhaus, S., Glöckner, G., & Valentin, K. (2012). A comprehensive cDNA library of light- and temperature-stressed *Saccharina latissima* (Phaeophyceae). *European Journal of Phycology, 47*, 83–94.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., & Vandesompele, J. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biology, 8*, R19.

Heo, S. J., Ko, S. C., Kang, S. M., Kang, H. S., Kim, J. P., Kim, S. H., … Jeon, Y. J. (2008). Cytoprotective effect of fucoxanthin isolated from brown algae *Sargassum siliquastrum* against H_2O_2-induced cell damage. *European Food Research and Technolology, 228*, 145-151.

Ho, C. L., Phang, S. M., Sinnappah, N. D., & Pang, T. (1996) Molecular approaches in the taxonomy of the red and brown seaweeds. In B.R. Chaudary, S.B. Agrawal (Eds.). *Cytology, Genetics and molecular biology of algae* (pp 351–362). Amsterdam: SPB Academic Publishing.

Ho, C. L., Kwan, Y. Y., Choi, M. C., Tee, S. S., Ng, W. H., Lim, K. A., …Meilina Ong Abdullah. (2007). Analysis and functional annotation of expressed sequence tags (ESTs) from multiple tissues of oil palm (*Elaeis guineensis* Jacq.). *BMC Genomics, 8*, 381.

Holdt, S. L. & Kraan, S. (2011). Bioactive compounds in seaweed: functional food applications and legislation. *Journal of Applied Phycology, 23*, 543-597.

Hong, Y. K., Sohn, C. H., Lee, K. W., & Kim, H. G. (1997). Nucleic acid extraction from seaweed tissues for polymerase chain reaction. *Journal of Marine Biotechnology, 5*, 95-99.

Hong, Y. K., Sohn, C. H., Polne-Fuller, M., & Gibor, A. (1995). Differential display of tissue-specific messenger RNAs in *Porphyra perforata* (Rhodophyta) thallus. *Journal of Phycology, 31*, 640-643.

Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., & Nakai, K. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Research, 35*, W585-W587.

Horváth, I., Glatz, A., Nakamoto, H., Mishkind, M. L., Munnik, T., Saidi, Y., …Vigh, L. (2012). Heat shock response in photosynthetic organisms: Membrane and lipid connections. *Progress in Lipid Research, 51*, 208-220.

Huang, C. Y., Wu, S. J., Yang, W. N., Kuan, A. W., & Chen, C. Y. (2016). Antioxidant activities of crude extracts of fucoidan extracted from *Sargassum glaucescens* by a compressional-puffing-hydrotheral extraction process. *Food Chemistry, 197*, 1121-1129.

Huang, X., & Madan, A. (1999). CAP3: A DNA sequence assembly program. *Genome Research, 9*, 868–877.

Japelaghi, R. H., Haddad, R., & Garoosi, G. A. (2011). Rapid and efficient isolation of high quality nucleic acids from plant tissues rich in polyphenols and polysaccharides. *Molecular Biotechnology, 49*, 129–137.

Jonsson, P. R., Granhag, L., Moschella, P. S., Åberg, P., Hawkins, S. J., & Thompson, R. C. (2006). Interactions between wave action and grazing control the distribution of intertidal macroalgae. *Ecology, 87*, 1169-1178.

Jothinayagi, N., & Anbazhagan, C. (2009). Heavy metal monitoring of Rameswaram Coast by some *Sargassum* species. *American-Eurasian Journal of Scientific Research, 4*, 73-80.

Kaliaperumal, N., Kalimuthu, S., & Ramalingam, J. R. (2004). Present scenario of seaweed exploitation and industry in India. *Seaweed Res. Utiln, 26* (1 & 2), 47-53.

Kamiya, M., Nishio, T., Yokoyama, A., Yatsuya, K., Nishigaki, T., Yoshikawa, S., & Ohki, K. (2010). Seasonal variation of phlorotannin in Sargassacean species from the coast of the Sea of Japan. *Phycological Research, 58*, 53-61.

Kanimozhi, A. S., Johnson, M., & Malar, R. J. J. T. (2015). Phytochemical composition of *Sargassum polycystum* C. Agardh and *Sargassum duplicatum* J. Agardh. *International Journal of Pharmacy and Pharmaceutical Sciences, 7*, 393-397.

Kanth, B. K., Kumari, S., Choi, S. H., Ha, H. J., & Lee, G. J. (2015). Generation and analysis of expressed sequence tags (ESTs) of *Camelina sativa* to mine drought stress-responsive genes. *Biochemical and Biophysical Research Communications, 467*, 83-93.

Karlenius, T. C., & Tonissen, K. F. (2010). Thioredoxin and cancer: a role for thioredoxin in all states of tumor oxygenation. *Cancers, 2*, 209-232.

Katz, L. A. (2012). Origin and diversification of eukaryotes. *Annual Review of Microbiology, 66*, 411-427.

Kenyon, F., Welsh, M., Parkinson, J., Whitton, C., Blaxter, M. L., & Knox, D. P. (2003). Expressed sequence tags survey of gene expression in the scab mite *Psoroptes ovis*-allergens, proteases and free-radical scavengers. *Parasitology, 126*, 451-460.

Killing, B., Cirik, S., Turan, G., Tekogul, H. & Koru, E. (2013). Seaweeds for food and industrial applications. In I. Muzzalupo (Ed.), *Food Industry*. InTech, Open access. doi: 10.5772/53172

Kim, K. M., Park, J. H., Bhattacharya, D., & Yoon, H. S. (2014). Applications of next-generation sequencing to unraveling the evolutionary history of algae. *International Journal of Systematic and Evolutionary Microbiology, 64*, 333-345.

Kim, S. H., Lee, Y. K., Hong, C. B., & Lee, I. K. (1997). A simple method for RNA extraction from marine macroalgae. *Algae, 12*, 53-56.

Komatsu, T., Fukuda, M., Mikami, A., Mizuno, S., Kantachumpoo, A., Tanoue, H., & Kawamiya, M. (2014). Possible change in distribution of seaweed, *Sargassum horneri*, in northeast Asia under A2 scenario of global warming and consequent effect on some fish. *Marine Pollution Bulletin, 85*, 317-324.

Kong, Q., Yuan, J., Niu, P., Xie, J., Jiang, W., Huang, Y., & Bie, Z. (2014). Screening suitable reference genes for normalization in reverse transcription quantitative real-time PCR analysis in melon. *PLoS ONE, 9*, e87197.

Kowalczyk, N., Rousvoal, S., Hervé, C., Boyen, C., & Collén, J. (2014). RT-qPCR normalization genes in the red alga *Chondrus crispus*. *PLoS One, 9*, e86574.

Kumar, M., Kumari, P., Gupta, V., Reddy, C. R. K., & Jha, B. (2010). Biochemical responses of red alga *Gracilaria corticata* (Gracilariales, Rhodophyta) to salinity induced oxidative stress. *Journal of Experimental Marine Biology and Ecology, 391*, 27-34.

Kumari, R., Kaur, I., & Bhatnagar, A. K. (2013). Enhancing soil health and productivity of *Lycopersicon esculentum* Mill. using *Sargassum johnstonii* Setchell & Gardner as a soil conditioner and fertilizer. *Journal of Applied Phycology, 25*, 1225-1235.

Küpper, F. C., Carpenter, L. J., McFiggans, G. B., Palmer, C. J., Waite, T. J., Boneberg E-M., … Feiters, M. C. (2008). Iodide accumulation provides kelp with an inorganic antioxidant impacting atmospheric chemistry. *Proceedings of the National Academy of Sciences, 105*, 6954-6958.

La Barre, S., Potin, P., Leblanc, C., & Delage, L. (2010). The halogenated metabolism of brown algae (Phaeophyta), its biological importance and its environmental significance. *Marine Drugs, 8*, 988-1010.

Leblanc, C., Colin, C., Cosse, A., Delage, L., La Barre, S., Morin, P., … Potin, P. (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, K. Y., & Mooney, D. J. (2012). Alginate: properties and biomedical applications. *Progress in Polymer Science, 37*, 106-126.

Lee, Y. M., Rhee, J. S., Hwang, D. S., Kim, I. C., Raisuddin, S. & Lee, J. S. (2007). Mining of biomarker genes from expressed sequence tags and differential display reverse transcriptase-polymerase chain reaction in the self-fertilizing fish, *Kryptolebias marmoratus* and their expression patterns in response to exposure to an endocrine-disrupting alkylphenol, bisphenol A. *Moleculars and Cells, 23*, 287-303.

Li, B., Lu, F., Wei, X., & Zhao, R. (2008). Fucoidan: structure and bioactivity. *Molecules, 13*, 1671-1695. doi: 10.3390/molecules13081671.

Liu, F. & Pang, S. J. (2010). Stress tolerance and antioxidant enzymatic activities in the metabolisms of the reactive oxygen species in two intertidal red algae *Grateloupia turuturu* and *Palmaria palmata*. *Journal of Experimental Marine Biology and Ecology, 382*, 82-87.

Liu, F., Sun, X., Wang, W., Liang, Z., & Wang, F. (2014). *De novo* transcriptome analysis-gained insights into physiological and metabolic characteristics of *Sargassum thunbergii* (Fucales, Phaeophyceae). *Journal of Applied Phycology, 26*, 1519-1526.

Liu, L., Heinrich, M., Myers, S., & Dworjanyn, A. A. (2012). Towards a better understanding of medicinal uses of the brown seaweed *Sargassum* in traditional chinese medicine: a phytochemical and pharmacological review. *Journal of Ethnopharmacology, 142*, 591-619.

Liu, J. J., Goh, C. J., Loh, C. S., Liu, P., & Pua, E. C. (1998). A method for isolation of total RNA from fruit tissues of banana. *Plant Molecular Biology Reporter, 16*, 1-6.

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

Lodeiro, P., Herrero, R., & Sastre de Vicente. (2006). The use of protonated *Sargassum muticum* as biosorbent for cadmium removal in a fixed-bed column. *Journal of Hazardous Materials, 137*, 244-253.

Lommer, M., Roy, A.S., Schilhabel, M., Schreiber, S., Rosenstiel, P., & LaRoche, J. (2010). Recent transfer of an iron-regulated gene from the plastid to the nuclear genome in an oceanic diatom adapted to chronic iron limitation. *BMC Genomics, 11*, 718.

Lorbeer, A. J., Lahnstein, J., Bulone, V., Nguyen, T., & Zhang, W. (2015). Multiple-response optimization of the acidic treatment of the brown alga *Ecklonia radiata* for the sequential extraction of fucoidan and alginate. *Bioresource Technology, 197*, 302-309.

Lyczak, J. B., Cannon, C. L., & Pier, G. B. (2002). Lung infections associated with cystic fibrosis. *Clinical Microbiology Reviews, 15*, 194-222.

MacKay R. M., & Gallant, J. W. (1991). Beta-tubulins are encoded by at least four genes in the brown alga *Ectocarpus variabilis*. *Plant Molecular Biology, 17*, 487-492.

Macovei, A., Vaid, N., Tula, S., & Tuteja, N. (2012). A new DEAD-box helicase ATP-binding protein (OsABP) from rice is responsive to abiotic stress. *Plant Signaling & Behaviour, 7*, 1138-1143.

Maiti, S., Singha, K., Ray, S., Dey, P., & Sa, B. (2009). Adipic acid dihydrazide treated partially oxidized alginate beads for sustained oral delivery of flurbiprofen. *Pharmaceutical Development and Technology, 14*, 461-470.

Marin, A., Casas-Valdez, M., Carrillo, S., Hernández, H., Monroy, A., Sanginés, L., & Pérez-Gil, F. (2009). The marine algae *Sargassum* spp. (Sargassaceae) as feed for sheep in tropical and subtropical regions. *Revista de Biologia Tropical, 57*, 1271-1281.

Market and Market. (February 2015). Alginates & derivatives market by type (sodium alginate, calcium alginate, potassium alginate, PGA, others), application (food & beverage, industrial, pharmaceuticals, others), & by region – global trends & forecast to 2019. Retrieved on 13 February, 2016 from http://www.marketsandmarkets.com/PressReleases/alginates-derivatives.asp

Marques, C. T., de Azevedo, T. C. G., Nascimento, M. S., Medeiros, V. P., Alves, L. G., Benevides, N. M. B., … Leite, E. L. (2012). Sulfated fucans extracted from algae *Padina gymnospora* have anti-inflammatory effect. *Brazilian Journal of Pharmacognosy, 22*, 115-122.

Marta, P., Ana, M. F., & Antonio, C. (2006). Biosorption of radioactive thorium by *Sargasum filipendula*. *Applied Biochemistry and Biotechnology, 134*, 193-206.

Martone, P. T., Alyono, M. & Stites, S. (2010). Bleaching of an intertidal coralline alga: untangling the effects of light, temperature and desiccation. *Marine Ecology Progress Series, 416*, 57-67.

Matloub, A. A., & Awad, N. (2012). Phytochemistry of some *Sargassum* spp. and their cytotoxic and antimicrobial activities. *Egyptian Pharmaceutical Journal, 11*, 99-108.

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

Mattio, L., & Payri, C. E. (2010). 190 years of *Sargassum* taxonomy, facing the advent of DNA phylogenies. *The Botanical Review, 77*, 31-70.

Mattio, L., Payri, C., & Stiger-Pouvreau, V. (2008). Taxonomic revision of *Sargassum* (Fucales, Phaeophyceae) from French Polynesia based on morphological and molecular analyses. *Journal of Phycology, 44*, 1541-1555.

Mattio, L., Payri, C., & Verlaque, M. (2009). Taxonomic revision and geographic distribution of subgen. *Sargassum* (Fucales, Phaeophyceae) in the western and central Pacific islands based on orphological and molecular analyses. *Journal of Phycology, 45*, 1213-1227.

Mattio, L., Payri, C. E., Verlaque, M., & de Reviers, B. (2010). Taxonomic revision of *Sargassum* sect. *Acanthocarpicae* (Fucales, Phaeophyceae). *Taxonomy, 59*, 896-904.

McHugh, D. J. (2003). A guide to seaweed industry. FAO Fisheries Technical Paper 441. Food and Agriculture Organization of the Inuted Nations (FAO) (p.105). Italy: Rome. Retrieved from http://www.fao.org/3/a-y4765e.pdf

McMillan, M., & Pereg, L. (2014). Evaluation of reference genes for gene expression analysis using quantitative RT-PCR in *Azospirillum brasilense*. *PLoS One, 9*, e98162.

Menon, V. V. (2012). Seaweed polysaccharides – food applications. In S.K. Kim (Ed.), *Handbook of marine macroalgae: Biotechnology and applied phycology* (pp. 541-555). United Kingdom, UK: John Wiley & Sons, Ltd.

Mekhalfi, M., Puppo, C., Avilan, L., Lebrun, R., Mansuelle, P., Maberly, S. C., & Gontero, B. (2014). Glyceraldehyde-3-phosphate dehydrogenase is regulated by

ferredoxin-NADP reductase in the diatom *Asterionella formosa*. *New Phytologist, 203*, 414-423.

Merchantt, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, J., Witman, G. B., … Grossman, A. R. (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions, *Science, 318*, 245-250.

Michel, G., Tonon, T., Scornet, D., Cock, J. M. & Kloareg, B. (2010a). Central and storage carbon metabolism of the brown alga *Ectocarpus siliculosus*: insights into the origin and evolution of storage carbohydrates in eukaryotes. *New Phytologist, 188*, 67-81.

Michel, G., Tonon, T., Scornet, D., Cock, J. M., & Kloareg, B. (2010b). The cell wall polysaccharide metabolism of the brown alga *Ectocarpus siliculosus*. Insight into the evolution of extracellular matrix polysaccharides in Eukaryotes. *New Phytologist, 188*, 82-97.

Mišurcová, L. (2012). Chemical composition of seaweeds. In S.K. Kim (Edi.), *Handbook of Marine Macroalgae: Biotechnology and applied phycology* (pp. 491-499). United Kingdom, UK: John Wiley & Sons, Ltd.

Mitchell, A., Chang, H.Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., …Finn, R. D. (2014). The InterPro protein families database: The classification resource after 15 years. *Nucleic Acids Research*, *43*, D213–D221.

Mohsin, S., Mahadevan, R., & Kurup, G. M. (2014). Free-redical-scavenging activity and antioxidant effect of ascophyllan from marine brown algae *Padina tetrastromatica*. *Biomedicine & Preventive Nutrition, 4*, 75-79.

Mouritsen, O. (2013). The science of seaweeds: Marine macroalgae benefit people culturally, industrially, nutritionally, and ecologically. *American Scientist*. Retrieved from http://www.americanscientist.org/issue/feature/the-science-of-seaweeds/1

Nagaraj, S. H., Gasser, R. B., & Ranganathan, S. (2006). A hitchhiker‘s guide to expressed sequence tag (EST) analysis. *Briefing in Bioinformatics, 8*, 6-21.

Nakamura, Y., Sasaki, N., Kobayashi, M., Ojima, N., Yasuike, M., Shigenobu, Y., … Ikeo, K. (2013). The first symbiont-free genome sequence of marine red alga, susabi-nori (*Pyropia yezoensis*), PLoS ONE, 8, e57122.

Narina, S. S., Buyyarapu, R., Kottapalli, K. R., Sartie, A. M., Mohamed I. Ali, Robert, A., … Scheffler, B. (2011). Generation and analysis of expressed sequence tags

(ESTs) for marker development in yam (*Dioscorea alata* L.). *BMC Genomics, 12*, 100.

New World Encyclopedia. (2015). Seaweed. Retrieved on 15 January 2016 from http://www.newworldencyclopedia.org/entry/Seaweed

Ng, W. S. (2013). *Taxonomy and phylogenetics of Malaysian Sargassum species (Fucales, Phaeophyceae)*. (Unpublished doctoral dissertation). University of Malaya, Kuala Lumpur.

Ngo, D. H., & Kim, S. K. (2013). Sulfated polysaccharides as bioactive agents from marine algae. *International Journal of Biological Macromolecules, 62*, 70-75.

Nguyen, P. D., Ho, C. L., Harikrishna, J. A., Wong, M. C. V. L., & Raha Abdul Rahim. (2006). Generation and analysis of expressed sequence tags from the mangrove plant, *Acanthus ebracteatus* Vahl. *Tree Genetics & Genomes, 2*, 196-201.

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.

Nolan, T., Hands, R. E. & Bustin, S. A. (2006). Quantification of mRNA using real-time RT-PCR. *Nature Protocol, 1*, 1559-1582.

Nunn, F. G., Burgess, S. T. G., Innocent, G., Nisbet, A. J., Bates, P., & Huntley, J. F. (2011). Development of a serodiagnostic test for sheep scab using recombinant protein Pso o 2. *Molecular and Cellular Probes, 25*, 212-218.

Nyvall, P., Corre, E., Boisset, C., Barbeyron, T., Rousvoal, S., Scornet, D., Kloareg, B., & Boyen, C. (2003). Characterization of mannuronan C-5-epimerase genes from the brown alga *Laminaria digitata*. *Plant Physiology, 133*, 726-735.

O'Brien, M. A., Costin, B. N., & Miles, M. F. (2012). Using genome-wide expression profiling to define gene networks relevant to the study of complex traits: from RNA integrity to network topology. *International Review of Neurobiology, 104*, 91-133.

OMRI (5 February, 2015). Technical Evaluation Report. Retrieved on 17 February, 2016 from https://www.ams.usda.gov/sites/default/files/media/Alginates%20TR%202015.pdf

Opeolu, B .O., Bamgbose, O., Arowolo, T. A., & Adetunji, M. T. (2010). Utilization of biomaterials as adsorbents for heavy metals‘ removal from aqueous matrices. *Scientific Research and Essays, 5*, 1780-1787.

Padilha, F. P., de Franca, F. P., & da Costa, A. C. A. (2004). The use of waste biomass of *Sargassum* sp. for the biosorption of copper from simulated semiconducter effluents. *Bioresource Technology, 96*, 1511-1517.

Pádua, D., Rocha, E., Gargiulo, D., & Ramos, A. A. (2015). Bioactive compounds from brown seaweeds: phloroglucinol, fucoxanthin and fucoidan as promising therapeutic agents against breast cancer. *Phytochemistry Letters, 14*, 91-98.

Pandithurai, M., & Murugesan, S. (2014). Free radical scavenging activity of methanolic extract of brown alga *Spatoglossum asperum*. *Journal of Chemical and Pharmaceutical Research, 6*, 128-132.

Parapurath, S. N., Bernard, H. E., Subramaniamc, D. M., & Ramamurthy, R. (2012). A dimensional investigation on seaweeds: their biomedical and industrial applications. In S.K. Kim (Ed.), *Handbook of marine macroalgae: Biotechnology and applied phycology* (pp. 532-540). United Kingdom, UK: John Wiley & Sons, Ltd.

Pareek, C. S., Smoczynski, R., & Tretyn, A. (2011). Sequencing technologies and genome sequencing. *Journal of Applied Genetics, 52*, 413-435.

Park, S. J., Huh, J. W., Kim, Y. H., Lee, S. R., Kim, S. H., Kim, S. U., … Chang, K. T. (2013). Selection of internal reference genes for normalization of quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis in the canine brain and other organs. *Molecular Biotechnology, 54*, 47-57.

Parkinson, J., & Blaxter, M. (2009). Expressed sequence tags: an overview. In J. Parkinson (Ed.), *Expressed Sequence Tags (ESTs)* (pp. 1-12). New York, NY: Humana Press.

Pearson, G., Lago-Leston, A., Valente, M., & Serrão, E. (2006). Simple and rapid RNA extraction from freeze-dried tissue of brown algae and seagrasses. *European Journal of Phycology, 41*, 97-104.

Pearson, G. A., Hoarau, G., Lago-Leston, A., Coyer, J. A., Kube, M., Reinhardt, R., … Olsen, J. L. (2010). An expressed sequence tag analysis of the intertidal brown seaweeds *Fucus serratus* (L.) and *F. vesiculosus* (L.) (Heterokontophyta, Phaeophyceae) in response to abiotic stressors. *Marine Biotechnology, 12*, 195–213.

Petersen, T. N., Brunak, S., von Heijne, G., & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods, 8*, 785-786.

Pfaffl, M.W. (2004). Quantification strategies in real-time PCR. In S.A. Bustin (Ed.), *A-Z of quantitative PCR* (pp. 87-112). La Jolla, California, USA: International University Line (IUL).

Pfaffl, M. W., Tichopad, A., Prgomet, C., & Neuvians, T. P. (2004). Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper—Excel-based tool using pair-wise correlations. *Biotechnology Letters, 26*, 509–515.

Phang, S. M., Wong, C. L., Ng, W. S., & Sim, M. C. (2008). Checklist of Malaysian *Sargassum* species. In S.M. Phang, K. Lewmanomont., & P.E. Lim (Eds.), *Taxonomy of Southeast Asian seaweeds: IOES Monograph Series 2* (pp. 83-103). Kuala Lumpur, Malaysia: Institute of Ocean and Earth Sciences (IOES), University of Malaya.

Phillips, N. E. Smith, C. M., & Morden, C. W. (2005). Testing systematic concepts of *Sargassum* (Fucales, Phaeophyceae) using portions of the *rbcLS* operon. *Phycological Research, 53*, 1-10.

Pielesz, A., & Biniaś, W. (2010). Cellulose acetate membrane electrophoresis and FTIR spectroscopy as methods of identifying a fucoidan in *Fucus vesiculosus* Linnaeus. *Carbohydrate Research, 345*, 2676-2682.

Pitsch, N. T., Witsch, B., & Baier, M. (2010). Comparison of the chloroplast peroxidase system in the chlorophyte *Chlamydomonas reinhardtii*, the bryophyte *Physcomitrella patens*, the lycophyte *Selaginella moellendorffii* and the seed plant *Arabidopsis thaliana*. *BMC Plant Biology, 10*, 133.

Pombert, J. F., & Keeling, P. J. (2010). The mitochondrial genome of the entomoparasitic green alga *Helicosporidium*. *PLoS ONE, 5*, e8954.

Pomin, V. H. (2012). Structure and use of algal sulfated fucans and galactans. In S.K. Kim (Ed.), *Handbook of marine macroalgae: Biotechnology and applied phycology* (229-261). United Kingdom, UK: John Wileys & Sons, Ltd.

Portune, K. J., Cary, S. C., & Warner, E. W. (2010). Antioxidant enzyme response and reactive oxygen species production in marine raphidophytes. *Journal of Phycology, 46*, 1161-1171.

Prabu, D. L., Sahu, N. P., Pal, A. K., & Narendra, A. (2013). Isolation and evaluation and antibacterial activities of fucoidan rich extract (FRE) from Indian brown seaweed. *Continental Journal of Pharmaceutical Sciences, 7*, 9-16.

Qi, M. (2014). Transplantation of encapsulated pancreatic islets as a treatment for patients with type I diabetes mellitus. *Advances in Medicine, 2014*, pp15. doi:10.1155/2014/429710

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., & Lopez, R. (2005). InterProScan: protein domains identifier. *Nucleic Acids Research, 33*, W116-W120.

Raghavendran, H. B., Sathivel, A., & Devaki, T. (2006). Defensive nature of *Sargassum polycystum* (Brown alga) against acetaminophen-induced toxic hepatitis in rats: Role of drug metabolizing microsomal enzyme system, tumor necrosis factor-α and fate of liver cell structural integrity. *World Journal of Gastroenterology, 12*, 3829-3834.

Rajauria, G., Cornish, L., Ometto, F., Msuya, F. E., & Villa, R. (2015). Identification and selection of algae for food, feed and fuel applications. In B.K. Tiwari & D. Troy (Eds.), *Seaweed sustainability, food and non-food applications* (pp. 315-345). Amsterdam: Academic Press, Elsevier Inc. doi:10.1016/B978-0-12-418697-2.00012-X

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

Ramakrishnan, P., Maclean, M., MacGregor, S. J., Anderson, J. G., & Grant, M. H. (2016). Cytotoxic responses to 405 nm light exposure in mammalian and bacterial cells: involvement of reactive oxygen species. *Toxicology In Vitro, 33*, 54-62.

Ramsey, D. M., & Wozniak, D. J. (2005). Understanding the control of *Pseudomonas aeruginosa* alginate synthesis and the prospects for management of chronic infections in cystic fibrosis. *Molecular Microbiology, 56*, 309-322.

Rao, P. V. S., & Mantri, V. A. (2006). Indian seaweed resources and sustainable utilization: Scenario at the dawn of a new century. *Current Science, 91*, 164-174.

Rattaya, S., Benjakul, S., & Prodpran, T. (2015). Extraction, antioxidative, and antimicrobial activities of brown seaweed extracts, *Turbinaria ornata* and *Sargassum polycystum*, grown in Thailand. *International Aquatic Research, 7*, 1-16.

Ren, Y., Perepelov, A. V., Wang, H., Zhang, H., Knirel, Y. A., Wang, L., & Chen, W. (2010). Biochemical characterization of GDP-L-fucose *de novo* synthesis pathway in fungus *Mortierella alpina*. *Biochemical and Biophysical Reseach Communications, 391*, 1663-1669.

Rio, D. C., Ares, M. Jr., Hannon, G. J., & Nilsen, T. W. (2010). Enrichment of poly(A)$^+$ mRNA using immobilized oligo(dT). *Cold Spring Harbor Protocols, 2010*, pdb.prot5454. doi: 10.1101/pdb.prot5454.

Ritter, A., Dittami, S. M., Goulitquer, S., Correa, J. A., Boyen, C., Potin, P., & Tonon, T. (2014). Trancriptomic and metabolomic analysis of copper stress acclimation in Ectocarpus siliculosus highlights signaling and tolerance mechanisms in brown algae. *BMC Plant Biology, 14*,116.

Ritter, A., Goulitquer, S., Salaün, J.P., Tonon, T., Correa, J.A.,& Potin, P. (2008). Copper stress induces biosynthesis of octadecanoid and eicosanoid oxygenated derivatives in the brown algal kelp *Laminaria digitata*. *New Phytologist, 180*, 809-821.

Roeder, V., Collén, J., Rousvoal, S., Corre, E., Leblanc, C., & Boyen, C. (2005). Identification of stress gene transcripts in *Laminaria digitata* (Phaeophyceae) protoplast cultures by expressed sequence tag analysis. *Journal of Phycology, 41*, 1227-1235.

Ronquist, F., Huelsenbeck, J., & Teslenko, M. (2011). MrBayes version 3.2 manual: Tutorials and model summaries. Retrieved from http://mrbayes.sourceforge.net/

Rousvoal, S., Groisillier, A., Dittami, S. M., Michel, G., Boyen, C., & Tonon, T. (2011). Mannitol-1-phosphate dehydrogenase activity in *Ectocarpus siliculosus*, a key role for mannitol synthesis in brown algae. *Planta, 233*, 261-273.

Rozen, S., & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology, 132*, 365–386.

Rudd, S. (2003). Expressed sequence tags: alternative or complement to whole genome sequnces? *Trends in Plant Science, 8*, 321-329.

Rupérez, P., Gómez-Ordóñez, E., & Jiménez-Escrig, A. (2014). Biological activity of algal sulfated and nonsulfated polysaccharides. In B. Hernandez-Ledesma & M. Herrero (Eds.), *Bioactive compounds from marine foods: plant and animal sources* (pp. 219-248). United Kingdom, UK: Wiley-Blackwell.

Sabra, W., & Zeng, A. P. (2009). Microbial production of alginates: Physiology and process aspects. In B.H.A. Rehm (Ed.), *Alginates: Biology and Applications*, Microbiology Monographs (Vol. 13) (pp. 153-173). Germany, Berlin, Heidelberg: Springer-Verlag. doi: 10.1007/978-3-540-92679-5_7.

Samaraweera, A. M., Vidanarachchi, J. K., & Kurukulasuriya, M. S. (2012). Industrial applications of macroalgae. In S.K. Kim (Ed.), *Handbook of marine macroalgae: Biotechnology and applied phycology* (pp. 500-521). United Kingdom, UK: John Wiley & Sons, Ltd.

Sambrook, J., & Russell, D. W. (2001). Molecular cloning: a laboratory manual (Vol. 2) (p. 11.63). New York, NY: Cold Spring Harbour Laboratory Press.

Santelices, B. (2007). The discovery of kelp forests in deep-water habitats of tropical regions. *Proceedings of the National Academy of Sciences, 104*, 19163-19164.

Saraswathi, S. J., Babu, B., & Rengasamy, R. (2003). Seasonal studies on the alginate and its biochemical composition I: *Sargassum polycystum* (Fucales), Phaeophyceae. *Phycological Research, 51*, 240-243.

Saravanan, A., Brindha, V., & Krishnan, S. (2011). Characteristic study of the marine algae *Sargassum* sp. on metal adsorption. *American Journal of Applied Sciences, 8*, 691-694.

Sargin, I., Arslan, G., & Kaya, M. (2016). Efficient of chitosan-algal biomass composite microbeads at heavy metal removal. *Reactive and Functional Polymers, 98*, 38-47.

Sartal, C. G., Alonso, M. C. B., & Barrera, P. B. (2012). Application of seaweeds in the food industry. In S.K. Kim (Ed.), *Handbook of marine macroalgae: Biotechnology and applied phycology* (pp. 522-531). United Kingdom, UK: John Wiley & Sons, Ltd.

Seenivasan, R., Rekha, M., Indu, H., & Geetha, S. (2012). Antibacterial activity and phytochemical analysis of selected seaweeds from mandapan coast, India. *Journal of Applied Pharmaceutical Science, 2*, 159-169.

Seidman, J. G. (2006). Chapter 5: Construction of recombinant DNA libraries. In F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, & K. Struhl (Eds.), *Current Protocols in Molecular Biology*. (pp. 5.8.7). Canada: John Wiley & Sons, Inc.

Shen, D., Sun, H., Huang, M., Zheng, Y., Qiu, Y., Li, X., & Fei, Z. (2013). Comprehensive analysis of expressed sequence tags from cultivated and wild radish (*Raphanus* spp.). *BMC Genomics, 14*, 721.

Shen, S., Zhang, G., Li, Y., Wang, L., Xu, P., & Yi, L. (2011). Comparison of RNA expression profiles on generations of *Porphyra yezoensis* (Rhodophyta), based on suppression substractive hybridization (SSH). *BMC Research Notes, 4*, 428.

Shiraya, T., Mori, T., Maruyama, T., Sasaki, M., Takamatsu, T., Oikawa, K., … Mitsui, T. (2015). Golgi/plastid-type manganese superoxide dismutase involved in heat-stress tolerance during grain filling of rice. *Plant Biotechnology Journal, 13*, 1251-1263.

Silberfeld, T., Leigh, J. W., Verbruggen, H., Cruaud, C., de Reviers, B., & Rousseau, F. (2010). A multi-locus time-calibrated phylogeny of the brown algae (Heterokonta, Ochrophyta, Phaeophyceae): Investigating the evolutionary nature of the ―brown algal crown radiation". *Molecular Phylogenetics and Evolution, 56*, 659-674.

Sim, M. C., Ho, C. L., & Phang, S. M. (2013). A simple and effective method for RNA isolation and cDNA library construction from the brown seaweed *Sargassum polycystum* (Fucales, Phaeophyceae). *Journal of Applied Phycology, 25*, 1277-1285.

Skjåk-Braek, G., & Draget, K. I. (2012). Alginates: properties and applications. In K. Matyjaszewski, & M. Möller. (Eds.), *Polymer Science: A Comprehensive Reference* (Vol. 10) (pp. 213-220). ScienceDirect: Elsevier B.V. doi:10.1016/B978-0-444-53349-4.00261-2

Skjåk-Braek, G., Larsen, B., & Grasdalen, H. (1986). Monomer sequence and acetylation pattern in some bacterial alginates. *Carbohydrate Research, 154*, 239-250.

Soanes, D. M., & Talbot, N. J. (2006). Comparative genomic analysis of phytopathogenic fungi using expressed sequence tag (EST) collections. Molecular *Plant Pathology, 7*, 61-70.

Stanley, M. S., Perry, R. M., Callow, J. A. (2005). Analysis of expressed sequence tags from the green alga *Ulva linza* (Chlorophyta). *Journal of Phycology, 41*, 1219-1226.

Steffy, J. A., Parveen, M. H., Durga, V., & Manibalan, S. (2013). Extraction purification of phlorotannins from different species of marine algae and

evaluation of their anti-oxidant potential. *Research Journal of Engineering and Technology, 4*, 163-168.

Sui, S., Luo, J., Ma, J., Zhu, Q., Lei, X., & Li, M. (2012). Generation and analysis of expressed sequence tags from *Chimonanthus praecox* (wintersweet) flowers for discovering stress-responsive and floral development-related genes. *Comparative and Functional Genomics, 2012*, Article ID 134596, 13 pages. http://dx.doi.org/10.1155/2012/134596

Tabaraki, R., & Nateghi, A. (2014). Multimetal biosorption modeling of Zn^{2+}, Cu^{2+} and Ni^{2+} by *Sargassum ilicifolium*. *Ecological Engineering, 71*, 197-205.

Tammam, A. A., Fakhry, E. M., & El-Sheekh, M. (2011). Effect of salt stress on antioxidant system and the metabolism of the reactive oxygen species on *Dunaliella salina* and *Dunaliella tertiolecta*. *African Journal of Biotechnology, 10*, 3795-3808.

Teo, S. S., Ho, C. L., Teoh, S. S., Lee, W. W., Tee, J. M., Raha Abdul Rahim & Phang, S. M. (2007). Analyses of expressed sequence tags from agarophyte, *Gracilaria changii* (Gracilariales, Rhodophyta). *European Journal of Phycology, 42*, 41-46.

Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2002). Multiple sequence aligment using ClustalW and ClustalX. *Current Protocols in Bioinformatics, 2.3:* 2.3.1-2.3.22.

Thomsen, J. & Melzner, F. (2010). Moderate seawater acidification does not elicit long-term metabolic depression in the blue mussel *Mytilus edulis*. *Marine Biology, 157*, 2667-2676.

Timperio, A. M., Egidi, M. G., & Zolla, L. (2008). Proteomics applied on plant abiotic stresses: Role of heat shock proteins (HSP). *Journal of Proteomics, 71*, 391-411.

Tipton, P. A. (2010). Synthesis of alginate in bacteria. In J. Reedijk, E. Dalcanale, B. Krebs, R. Marquardt, M. Morbidelli, H. Nakai, L. Panza, C. Poole, M. Quack, & K. Wandelt (Eds.), *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering: Comprehensive Natural Products II* (Vol. 8) (pp. 423-441). ScienceDirect: Elsevier B. V. doi: 10.1016/B978-008045382-8.00651-1

Tonon, T., Rousvoal, S., Roeder, V., & Boyen, C. (2008). Expression profiling of the mannuronan C5-epimerase multigenic family in the brown alga *Laminaria digitata* (Phaeophyceae) under biotic stress conditions. *Journal of Phycology, 44*, 1250-1256.

Urzica, E. I., Adler, L. N., Page, M. D., Linster, C. L., Arbing, M. A., Casero, D., … Clarke, S. G. (2012). Impact of oxidative stress on ascorbate biosynthesis in *Chlamydomonas* via regulation of the VTC2 gene encoding a GDP-L-galactose phosphorylase. *The Journal of Biological Chemistry, 287*, 14234-14245.

Vandesompele, J., De Preter, K., Poppe, B., Van Roy, N., De Paepe, A. & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology, 3*, 0034.1-0034.11.

Vavilala, S. L., & D'Souza, S. (2015). Algal polysaccharides and their biological applications. In S.K. Kim & K. Chojnacka (Eds.), *Marine algae extracts: processes, products, and applications* (pp. 413-452). United Kingdom, UK: John Wiley & Sons, Ltd.

Vellosillo, T., Aguilera, V., Marcos, R., Bartsch, M., Vicente, J., Cascón, T., Hamberg, M., & Castresana, C. (2013). Defense activated by 9-lipoxygenase-derived oxylipins requires specific mitochondrial proteins. *Plant Physiology, 161*, 617–627.

Victor, B., Dorny, P., Kanobana, K., Polman, K., Lindh, J., Deelder, A. M., … Gabriël, S. (2013). Use of expressed sequence tags as an alternative approach for the identification of *Taenia solium* metacestode excretion/ secretion proteins. *BMC Research Notes, 6*, 224.

Wahl, M., Jormalainen, V., Eriksson, B. K., Coyer, J. A., Molis, M., Schubert, H., … Olsen, J. L. (2011). Stress ecology in *Fucus*. Abiotic, biotic and genetic interactions. *Advances in Marine Biology, 59*, 37-105.

Wakasugi, T., Nagai, T., Kapoor, M., Sugita, M., Ito, M., Tsudzuki, J., … Sugiura, M. (1997). Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division. *Proceedings of National Academic of Sciences, 94*, 5967-5972.

Wang, G., Sun, J., Liu, G., Wang, L., Yu, J., Liu, T., … Wu, S. (2014a). Comparative analysis on transcriptome sequencings of six *Sargassum* species in China. *Acta Oceanologica Sinica, 33*, 37-44.

Wang, R., Wang, X., Zhang, Y., Yu, J., Liu, T., Chen, S., & Chi, S. (2014b). Origin and evolution of alginate-c5-mannuronan-epimerase gene based on transcriptomic analysis of brown algae. *Acta Oceanologica Sinica, 33*, 73–85.

Wang, W. J., Wang, F. J., Sun, X. T., Liu, F. L., Liang, Z. R. (2012). Comparison of transcriptome under red and blue light culture of *Saccharina japonica* (Phaeophyceae). *Planta, 237*, 1123-1133.

Wang, X., Tian, W., & Li, Y. (2008). Development of an efficient protocol of RNA isolation from recalcitrant tree tissues. *Molecular Biotechnology, 38*, 57-64.

Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., … Lander, E. S. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature, 420*, 520-562.

White, W. L., & Wilson, P. (2015). World seaweed utilization. In B.K. Tiwari & D. Troy (Eds.), *Seaweed sustainability: Food and Non-Food Applications* (pp. 7-25). Elsevier: Academic Press. doi:10.1016/B978-0-12-418697-2.00002-7.

Wichachucherd, B., Liddle, L. B., & Prathep, A. (2009). Population structure, recruitment, and succession of the brown alga, *Padina boryana* Thivy (Dictyotales, Heterokontophyta), at an exposed shore of Sirinart National Park and a sheltered area of Tang Khen Bay, Phuket Province, Thailand. *Aquatic Botany, 92*, 93-98.

Wiencke, C., & Amsler, C. D. (2012). Seaweeds and their communities in polar regions. In C. Wiencke & K. Bischof (Eds.), *Seaweed Biology, Ecological Studies* (pp. 265-291). Berlin, Heidelberg: Springer-Verlag.

Wijesinghe, W. A. J. P., & Jeon, Y. (2012). Biological activities and potential industrial applications of fucose rich sulfated polysaccharides and fucoidans isolated from brown seaweeds: a review. *Carbohydrate Polymers, 88*, 13-20.

Wikipedia. (2016a). Functional genomics. Retrieved on 12 February 2016 from https://en.wikipedia.org/wiki/Functional_genomics

Wikipedia. (2016b). *Macrocystid pyrifera*. Retrieved on 7 March 2016 from https://en.wikipedia.org/wiki/Macrocystis_pyrifera

Wilkins, T. A., & Smart, L. B. (1996). Isolation of RNA from plant tissue. In P.A. Kreig (Ed.), *A laboratory guide to RNA: isolation, analysis and synthesis* (pp. 21 - 41). New York, NY: Wiley-Liss, Inc.

Womersley H. B. S. (1954). Australian species of *Sargassum*, subgenus *Phyllotrichia*. *Australian Journal of Botany, 2*, 337-354.

Wong, C. L., Gan, S. Y., & Phang, S. M. (2004). Morphological and molecular characterization and differentiation of *Sargassum baccularia* and *S. polycystum* (Phaeophyta). *Journal of Applied Phycology, 16*, 439-445.

Wong, C. L., & Phang, S. M. (2004). Biomass production of two *Sargassum* species at Cape Rachado, Malaysia. *Hydrobiologia, 512*, 79–88.

Wong, T. K. M., Ho, C. L., Lee, W. W., Raha Abdul Rahim, & Phang, S. M. (2007). Analyses of expressed sequence tags from *Sargassum binderi* (Phaeophyta). *Journal of Phycology, 43*, 528-534.

WoundSource. (2016). Retrived on 3 January 2016 from http://www.woundsource.com/product-category/dressings/alginates

Xuan, J., Feng, Y., Weng, M., Zhao, G., Shi, J., Yao, J., … Wang, B. (2012). Expressed sequence tag analysis and cloning of trehalose-6-phosphate synthase gene from marine alga *Laminaria japonica* (Phaeophyta). *Acta Oceanologica Sinica, 31*, 139-148.

Yamagishi, T, Müller, D. G., & Kawai, H. (2014). Comparative transcriptome analysis of *Discosporangium mesarthrocarpum* (Phaeophyceae), *Schizocladia ischiensis* (Schizocladiophyceae), and *Phaeothamnion confervicola* (Phaeothamniophyceae), with special reference to cell wall-related genes. *Journal of Phycology*, *50*, 543-551.

Yang, L., Luo, Y., Wei, J., Ren, C., Zhou, X., & He, S. (2010). Methods for protein identification using expressed sequence tags and peptide mass fingerprinting for seed crops without complete genome sequences. *Seed Science Research, 20*, 257-262.

Yao, J., Fu, W., Wang, X., & Duan, D. (2009). Improved RNA isolation from *Laminaria japonica* Aresch (Laminariaceae, Phaeophyta). *Journal of Applied Phycology, 21*, 233-238.

Ye, N., Zhang, X., Miao, M., Fan, X., Zheng, Y., Xu, D., ... Zhao, F. (2015). *Saccharina* genomes provide novel insight into kelp biology. *Nature Communications, 6*, 6986.

Yeong, B. M. L., & Wong, C. L. (2013). Seasonal growth rate of *Sargassum* species at Teluk Kemang, Port Dickson, Malaysia. *Journal of Applied Phycology, 25*, 805-814.

Yockteng, R., Almeida, A. M. R., Yee, S., Andre, T., Hill, C., & Specht, C. D. (2013). A method for extracting high-quality RNA from diverse plants for next-generation sequencing and gene expression analyses. *Applications in Plant Sciences, 1*, 1300070.

Yoshida T. (1983). Japanese species of *Sargassum* subgenus *Bactrophycus* (Phaeophyta, Fucales). *Journal of the Faculty of Science, Hokkaido University, Series V (Botany), 13*, 99-246.

Yoshida, T. (1989). Taxonomy of *Sargassum*. *ALGAE, 4*, 107-110.

Yoshida, T., Ajisaka, T., Noro, T., & Horiguchi, T. (2004). Species of the genus *Sargassum* subgenus *Schizophycus*. In I.A. Abbott (Ed.), *Taxonomy of economic seaweeds* (Vol. 9) (pp. 93-106). La Jolla, California: California Sea Grant College, University of California.

Yu, Z., Zhu, X., Jiang, Y., Luo, P., & Hu, C. (2014). Bioremediation and fodder potentials of two *Sargassum* spp. in coastal waters of Shenzhen, South China. In R. Wu, B. Richardson, P. Shin, D. Au, & Kenneth (Eds.), *Marine Pollution Bulletin, 7th International Conference on Marine Pollution and Ecotoxicology, 85*, 797-802.

Zhang, X., Ye, N., Liang, C., Mou, S., Fan, X., Xu, J., … Zhuang, Z. (2012). *De novo* sequencing and analysis of the *Ulva linza* transcriptome to discover putative mechanisms associated with its successful colonization of coastal ecosystems. *BMC Genomics, 13*, 565.

Zhu, J., Zhang, L., Li, W., Han, S., Yang, W.,& Qi, L. (2013). Reference gene selection for quantitative real-time PCR normalization in *Caragana intermedia* under different abiotic stress conditions. *PLoS One, 8*, e53196.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

Sim, M. C., Ho, C. L., & Phang, S. M. (2013). A simple and effective method for RNA isolation and cDNA library construction from the brown seaweed *Sargassum polycystum* (Fucales, Phaeophyceae). *Journal of Applied Phycology, 25*(5), 1277-1285.

Phang, S. M., Wong, C. L., Ng, W. S., & Sim, M. C. (2008). Checklist of Malaysian *Sargassum* species. In S.M. Phang, K. Lewmanomont, & P.E. Lim (Eds.), Taxonomy of Southeast Asian Seaweeds, IOES Monograph Series 2 (pp. 83-103). Kuala Lumpur, Malaysia: Institute of Ocean and Earth Sciences, University of Malaya.

Sim, M. C., Chan, C. X., Ho, C. L., & Phang, S. M. (2016). Generation and analyses of expressed sequence tags (ESTs) from the brown macroalga *Sargassum polycystum* C. Agardh (Fucales, Ochrophyta). PLoS ONE. To be submitted.

Generation and analyses of expressed sequence tags (ESTs) from *Sargassum polycystum*, a brown seaweed. 7th Asia-Pacific Conference on Algal Biotechnology (APCAB), 1 - 4 December 2009, Department of Botany, University of Delhi, New Delhi, India.

Generation and analyses of expressed sequence tags (ESTs) from *Sargassum polycystum* (Fucales, Ochrophyta). Sharing Knowledge, Resources and Technologies for a Sustainable South China Sea 2012, 21 – 24 October 2012, Institute of Ocean and Earth Sciences, University of Malaya, Kuala Lumpur, Malaysia.

APPENDICES

Appendix A: Formula for Media and Solutions

Ampicilin (50 mg/mL)	
Ampicilin sodium salt	1g
dH_2O	20 mL
Filter sterilized through a 0.22 μm membrane	
Stored at -20 °C	
Kanamycin (50 mg/mL)	
Kanamycin	1 g
dH_2O	20 mL
Filter sterilized through a 0.22 μm membrane	
Stored at -20 °C	
Tetracycline (5 mg/ mL)	
Tetracycline	0.1 g
Absolute ethanol	20 mL
Stored in light-tight container at -20 °C	
0.5 M IPTG	
Dissolve 1.19 g in 10 mL sterile dH_2O	
Filter sterilized through 0.22 μm membrane	
X-Gal (250 mg/mL)	
Dissolve 1 g X-Gal in 4 mL of dimethylformamide	
Stored at -20 °C and protected from light	
LB Broth (1 L)	
NaCl	10 g
Tryptone	10 g
Yeast extract	5 g
Adjust to pH 7.0 and autoclaved	
LB Agar (1 L)	
NaCl	10 g
Yeast extract	5 g
Tryptone	10 g
Adjust to pH 7.0	
Agar	20 g
Autoclave	
NZY Agar (1 L)	
NaCl	5 g
$MgSO_4.7H_2O$	2 g
Yeast extract	5 g
NZ amine	10 g
Adjust to pH 7.5	

Agar	15 g
Autoclave	
NZY Top Agar (1 L)	
NaCl	5 g
$MgSO_4.7H_2O$	2 g
Yeast extract	5 g
NZ amine	10 g
Adjust to pH 7.5	
Agarose	7 g
Autoclave	
SM Buffer (1 L)	
NaCl	5.8 g
$MgSO_4.7H_2O$	2.0 g
1 M Tris-HCl (pH 7.5)	50 mL
2 % (w/v) gelatin solution	5 mL
Autoclave	
10 % Maltose (100 mL)	
Maltose	10 g
Filter sterilized through a 0.22 μm membrane	
TE Buffer	
10 mM Tris-HCl	
1 mM EDTA (pH 8.0)	
Autoclave	
TAE (Tris/acetate/EDTA) Buffer (50X) (1 L)	
Tris base	242 g
Glacial acetic acid	57.1 mL
0.5 M EDTA (pH 8.0)	100 mL

Appendix B: Malaysian Meteorological Department (MMD)

JABATAN METEOROLOGI MALAYSIA

Records of Daily Rainfall Amount (0800- 0800 MST)

Station : RUMAH API TANJUNG TUAN

Lat. : 2° 24´ N **Year : 2010**

Long. : 101° 51´ E **Unit : millimetre**

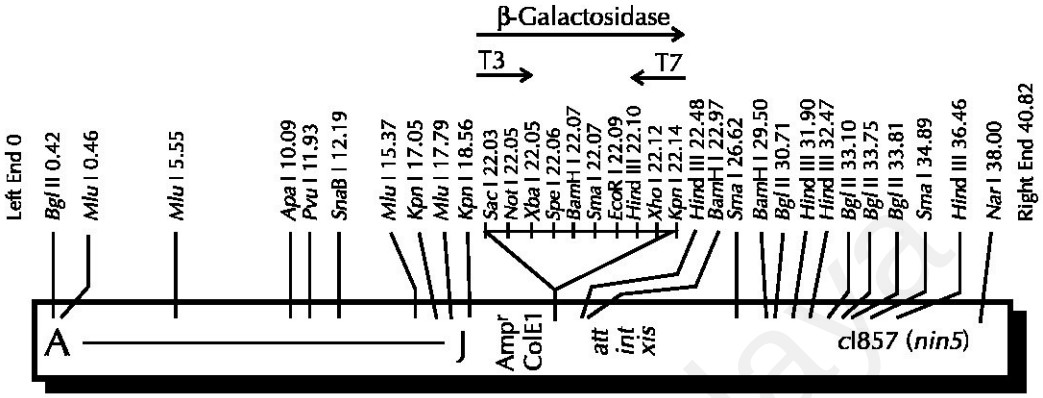
Ht. above M.S.L. : 0.0 m

Month/ Date	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
01	0.9	0.0	0.0	17.0	0.0	0.0	22.0	0.0	0.0	16.7	20.0	0.0
02	0.4	0.0	0.0	2.8	0.0	5.2	15.0	0.0	0.0	1.5	0.0	0.0
03	13.3	0.7	0.0	1.2	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
04	0.0	24.2	0.0	0.0	0.0	4.9	34.8	0.0	0.0	0.0	0.0	0.0
05	0.0	0.0	3.8	0.0	1.2	6.0	0.0	0.0	0.0	0.6	0.0	0.0
06	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	16.8	0.0	17.0	0.0
07	0.0	0.0	17.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	16.0	0.0
08	0.0	0.0	0.0	0.0	0.0	0.0	3.9	0.0	21.9	0.0	0.3	6.3
09	0.2	1.0	10.2	0.0	62.2	0.0	0.0	0.0	0.0	0.0	0.0	2.2
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	41.0	0.0	2.8	0.0
11	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	5.8	8.5	0.0
12	0.6	0.0	0.0	0.0	7.0	0.0	1.8	0.0	0.0	0.2	0.0	1.9
13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
14	0.0	0.0	0.0	0.0	31.0	0.0	3.9	0.0	0.0	0.0	4.5	0.0
15	0.0	0.0	0.0	0.0	0.0	0.0	14.1	0.0	0.0	0.0	0.0	1.5
16	0.0	0.0	0.0	43.0	9.0	0.0	27.1	6.0	0.0	0.0	2.0	0.0
17	0.0	0.0	0.0	2.2	0.0	0.0	0.0	0.0	0.0	0.0	7.0	0.0
18	14.4	18.3	0.0	4.7	6.0	0.0	6.3	0.0	0.0	0.0	0.0	0.0
19	0.0	0.0	0.0	7.1	0.0	0.0	0.0	3.0	0.0	0.0	0.0	0.0
20	2.8	0.0	0.0	0.0	3.0	0.0	0.0	0.0	5.2	5.2	0.0	1.2
21	0.0	0.0	7.8	24.3	2.8	12.2	11.2	0.0	0.0	0.0	0.0	0.0
22	0.0	0.0	1.2	0.0	26.0	9.3	0.0	0.0	5.8	3.7	0.0	24.8
23	0.0	4.0	0.0	67.4	0.0	1.2	0.0	0.0	28.2	7.8	0.0	2.8
24	0.0	2.1	0.0	0.0	0.0	65.5	0.0	4.0	1.3	2.3	17.3	0.0
25	0.0	0.0	1.5	0.0	0.0	3.2	0.0	0.0	2.2	0.8	0.0	0.0
26	0.0	0.0	0.0	6.5	4.0	49.1	0.0	0.0	1.2	0.5	0.0	0.0
27	0.0	0.0	0.0	0.0	0.0	0.0	1.2	8.8	0.0	0.4	0.0	0.0
28	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	30.1
29	0.0		1.0	0.0	27.0	0.0	3.3	0.0	0.0	0.0	3.7	3.0
30	0.0		0.0	0.0	0.0	7.2	7.2	0.0	31.8	5.3	0.8	0.0
31	0.0		0.0		0.0		0.0	0.0		0.0		0.0
Total	32.6	50.3	42.5	176.2	182.4	164.8	154.6	21.8	155.4	50.8	99.9	73.8
No. of Raindays	7	6	7	10	12	11	15	4	10	13	12	9

Definition : Trace - Rainfall amount less than 0.1 millimetre
Note : N.A. - Not Available
} - Accumulated Value

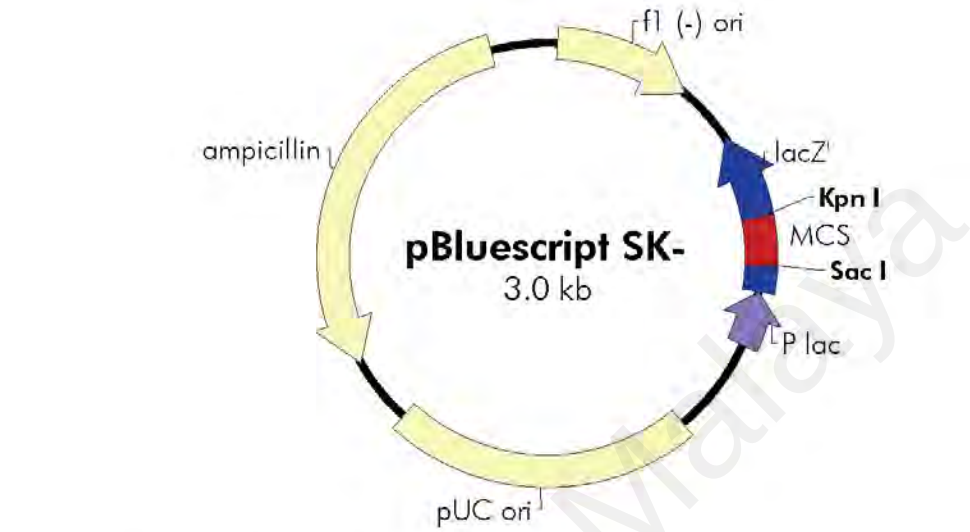
Dibekalkan oleh
Seksyen Kajiklim & Hidrologi
Jabatan Meteorologi Malaysia

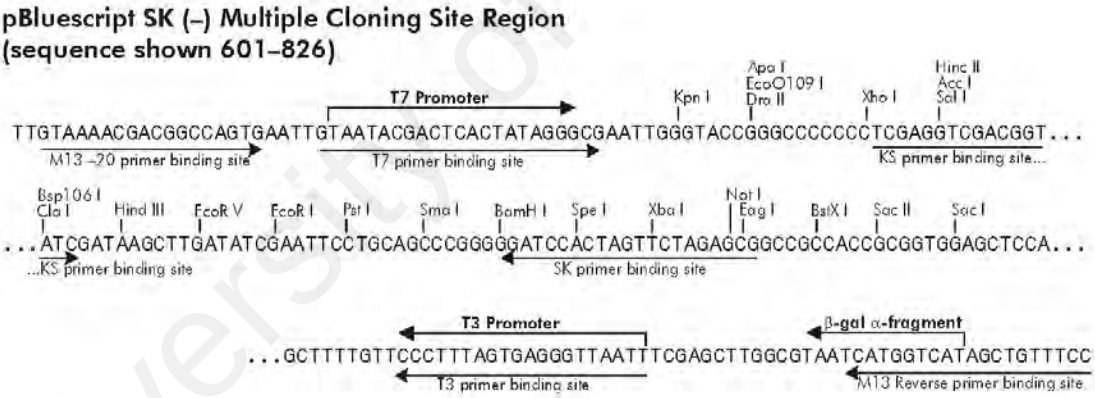
Appendix C: Uni-ZAP® XR Vector Map



Source: Stratagene cDNA Synthesis Kit, ZAP-cDNA® Synthesis Kit, and ZAP-cDNA® Gigapack® III Gold Cloning Kit Instruction Manual

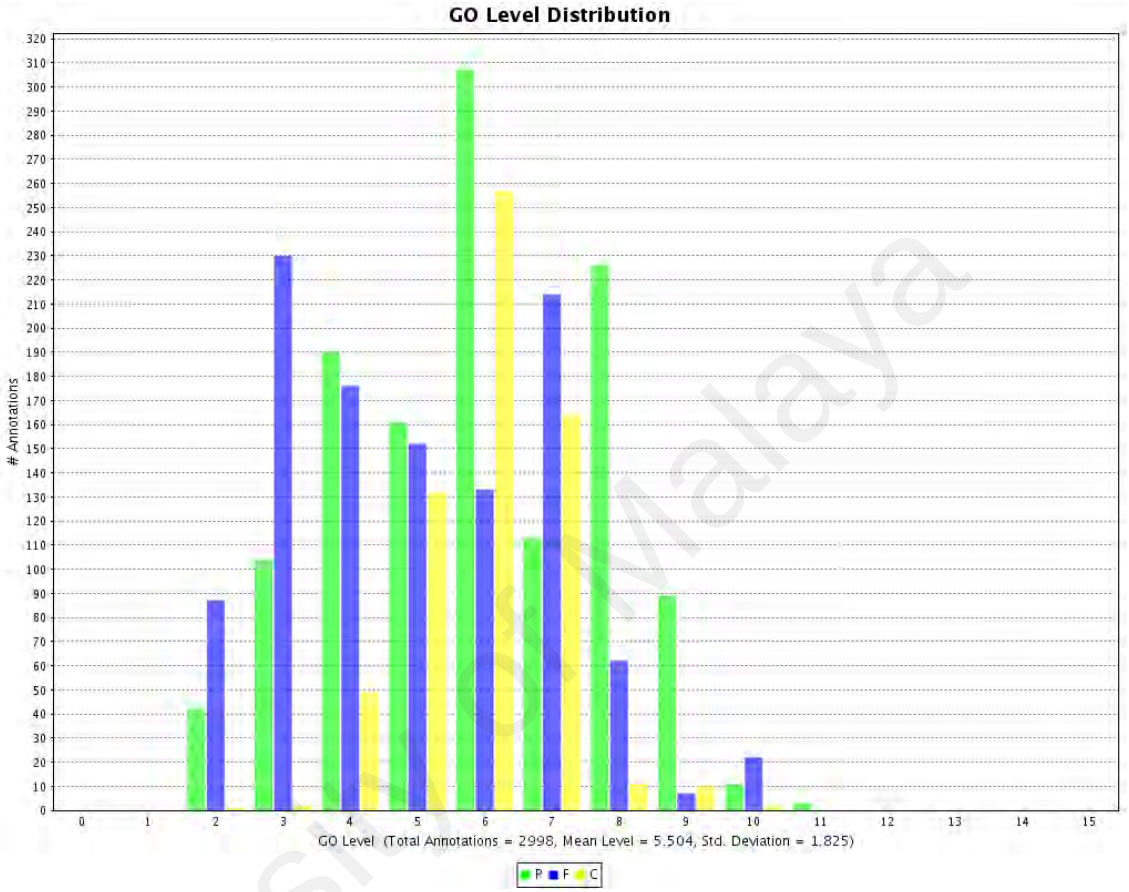
Appendix D: Vector map and polylinker sequence of the pBluescript SK(-) phagemid





Source: Stratagene cDNA Synthesis Kit, ZAP-cDNA® Synthesis Kit, and Zap-cDNA® Gigapack® III Gold Cloning Kit Instruction Manual

Appendix E: GO level distribution

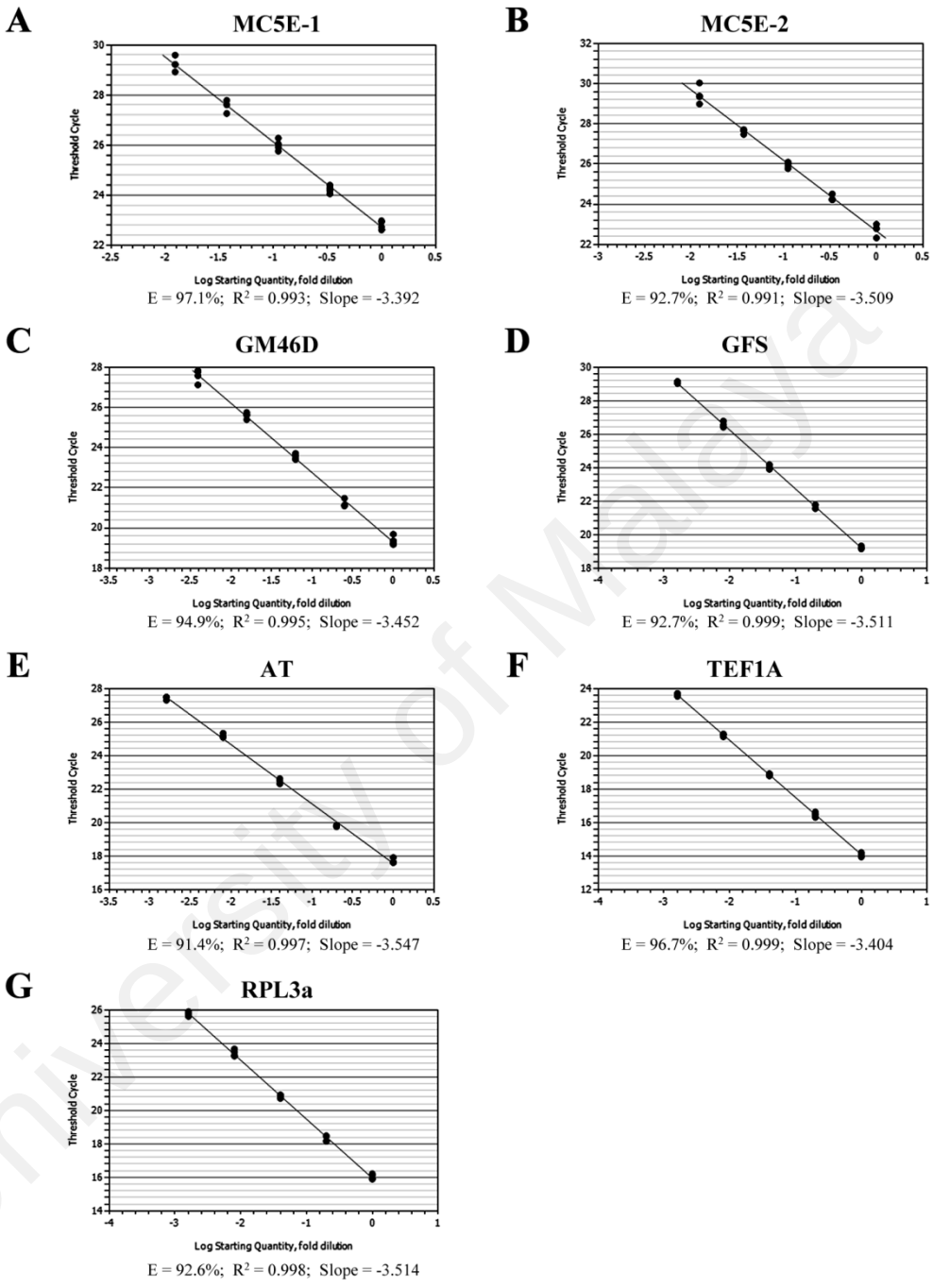


Of the 1142 unigenes with significant matches to NCBI nr protein database, 695 (~40.3 %) were functionally classified into one or more GO categories, resulting in a total of 2998 GO terms.

Appendix F: Categories of KEGG Pathways

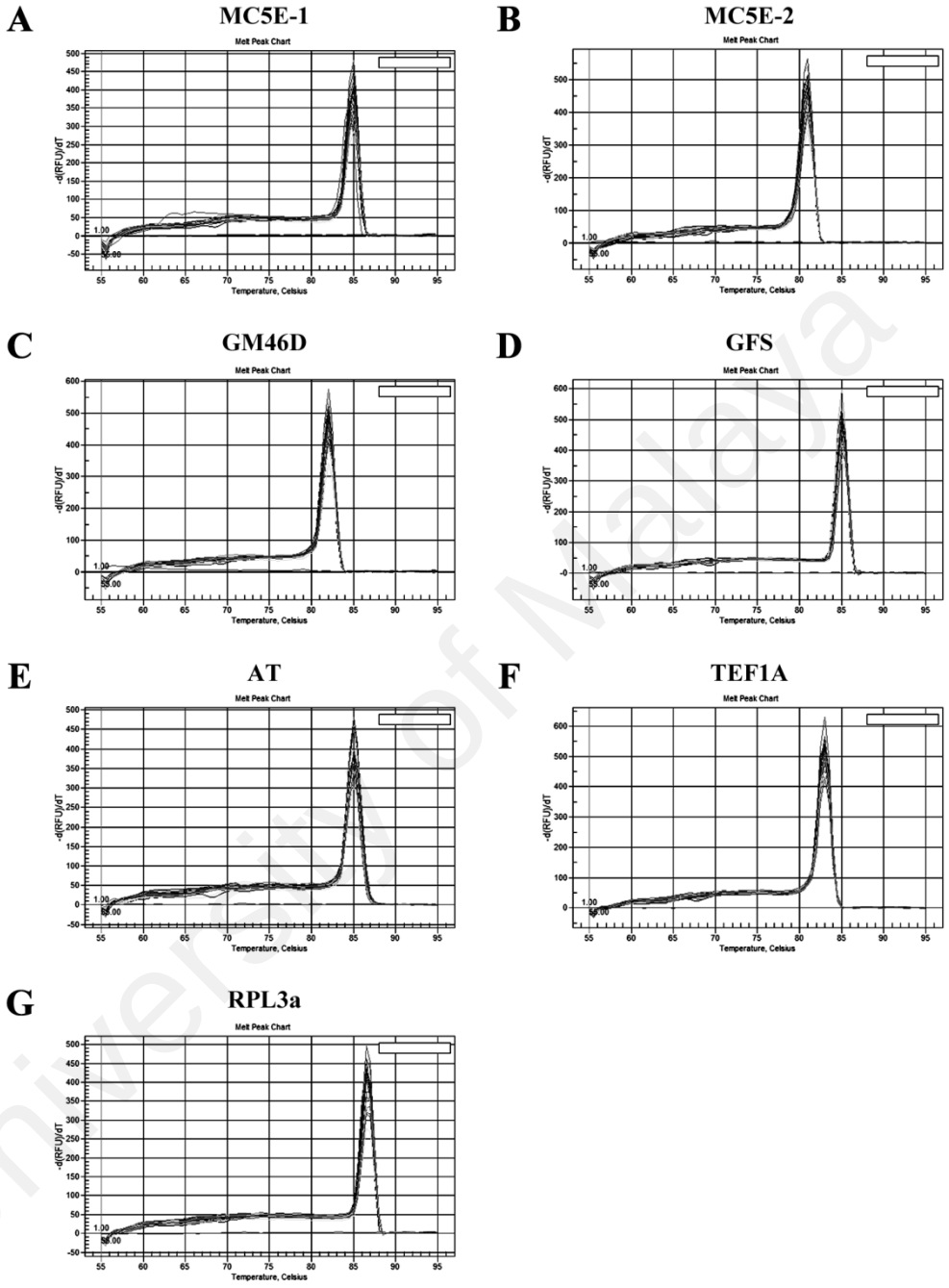
		Number of Pathways	Total unigenes in pathway
METABOLISM			
1.0	Global and overview map	1	43
1.1	Carbohydrate metabolism	13	63
1.2	Energy metabolism	6	34
1.3	Lipid metabolism	9	15
1.4	Nucleotide metabolism	2	12
1.5	Amino acid metabolism	11	33
1.6	Metabolism of other amino acids	2	7
1.8	Metabolism of cofactors and vitamins	4	10
1.9	Metabolism of terpenoids and polyketides	4	7
1.10	Biosynthesis of other secondary metabolites	4	4
1.11	Xenobiotics biodegradation and metabolism	5	8
GENETIC INFORMATION PROCESSING			
2.2	Translation	1	4
ENVIRONMENTAL INFORMATION PROCESSING			
3.2	Signal transduction	1	1
ORGANISMAL SYSTEMS			
5.1	Immune system	1	1

Appendix G: Standard Curve of RT-qPCR



Standard curves for mannuronan C5 epimerase MC5E-1 (A) and MC5E-2 (B), GDP-D-mannose 4,6 dehyratase (GM46D) (C), GDP-4-keto-6-D-mannose-epimerase-4-reductase (GFS) (D), Alpha Tubulin (AT) (E), Translation elongation factor 1 alpha (TEF1A) (F), Ribosomal protein L3a (RPL3a) (G).

Appendix H: Dissociation curves of genes obtained from RT-qPCR



Dissociation curves for mannuronan C5 epimerase MC5E-1 (A) and MC5E-2 (B), GDP-D-mannose 4,6 dehyratase (GM46D) (C), GDP-4-keto-6-D-mannose-epimerase-4-reductase (GFS) (D), Alpha Tubulin (AT) (E), Translation elongation factor 1 alpha (TEF1A) (F), Ribosomal protein L3a (RPL3a) (G).