UNIVERSITI PUTRA MALAYSIA

OPTIMIZATION OF RNA EXTRACTION AND GENERATION OF EXPRESSED SEQUENCE TAGS FROM SARGASSUM BINDERI (SONDER) J. AGARDH

TONY WONG KOK MIN.

FBSB 2005 15

OPTIMIZATION OF RNA EXTRACTION AND GENERATION OF EXPRESSED SEQUENCE TAGS FROM *Sargassum binderi* (SONDER) J. AGARDH

By

TONY WONG KOK MIN

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science

July 2005

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master Science

OPTIMIZATION OF RNA EXTRACTION AND GENERATION OF EXPRESSED SEQUENCE TAGS FROM *Sargassum binderi* (SONDER) J. AGARDH

By

TONY WONG KOK MIN

July 2005

Chairman : Ho Chai Ling, PhD

Faculty : Biotechnology and Biomolecular Sciences

Brown seaweeds are macroalgae in the marine habitat. They are widely used as food, sources for cosmetics, pharmaceuticals, soap, agar, textile, and alginate. *Sargassum binderi* is one of the most abundant brown seaweeds in Malaysia. Previous studies have shown that *S. binderi* contains high quality of alginate compared to commercially used seaweeds such as *Laminaria hyberborea*, *Macrocystis pyrifera* and *Ascophyllum nodosum*. Expressed sequence tag (EST) approach is a powerful tool in providing genetic information of an organism, especially for *S. benderi* in which little genetic information is available. This study presents the first attempt in generating ESTs from *S. binderi*. RNA extraction from seaweeds was the main challenge in this study, as the RNA yield was low and the polysaccharide contamination was difficult to be eliminated. A total of ten different RNA extraction methods (including five modifications) have been carried out to obtain sufficient RNA of high quality to construct a representative cDNA library for ESTs generation. The optimized *S. binderi*-

specific CTAB RNA extraction method developed in this study was able to produce high yield of RNA with minimum polysaccharide contamination. Sufficient amount of mRNA was obtained to construct a primary cDNA library with a titer of 9.2 x 10^5 pfu/ml. A total of 2051 ESTs were generated and analyzed from the amplified cDNA library of *S. binderi* (with a titer of 1.31 x 10^9 pfu/ml). The ESTs were putatively identified by comparison to the non-redundant peptide database in NCBI. Approximately 82% of the ESTs were assigned as unknown and novel sequences that are potentially important for new gene discovery, whereas the 18% ESTs with significant matches to the database were classified into various putative functional groups, including protein synthesis, energy, protein destination and storage, metabolism, cell structure/division, disease/defense, signal transduction, transcription, and miscellaneous. The EST information generated from this study may contribute towards better understanding of the biochemistry and molecular biology of *S. binderi* and other brown seaweeds in the future.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains

OPTIMASI PNGEKSTRAKAN RNA DAN PENJANAAN TAG JUJUKAN TEREKSPRES DARIPADA *Sargassum binderi* (SONDER) J. AGARDH

Oleh

TONY WONG KOK MIN

Julai 2005

Pengerusi : Ho Chai Ling, PhD

Fakulti : Bioteknologi dan Sains Biomolekul

Rumpai laut perang adalah makroalga di laut. Ia digunakan secara meluas di dalam industri makanan, kosmetik, farmasi, sabun, agar, tekstil and alginate. *Sargassum binderi* merupakan salah satu rumpai laut perang yang boleh didapati dengan banyaknya di Malaysia. Hasil penyelidikan awal menunjukkan bahawa *S. binderi* mengandungi alginate berkualiti tinggi berbanding dengan rumpai laut perang lain yang digunakan di dalam industri seperti *Laminaria hyberborea*, *Macrocystis pyrifera* dan *Ascophyllum nodosum*. Pendekatan tag jujukan terekspres (expressed sequence tag – EST) amat berguna untuk memperolehi maklumat genetik sesuatu organisma, terutamanya untuk *S. binderi* yang mempunyai maklumat genetik yang terhad. Kajian ini merupakan penyelidikan terulung untuk memperolehi EST daripada *S. binderi*. Pengekstrakan RNA merupakan suatu cabaran utama dalam pengajian ini, di mana kuantiti RNA yang diperolehi adalah rendah dan dicemari polisakarida yang susah dipisahkan. Sejumlah sepuluh kaedah pengekstrakan RNA (termasuk lima pengubahsuaian) telah

dijalankan untuk mendapat RNA yang cukup dan berkualiti demi pembinaan perpustakaan cDNA untuk penjanaan EST. Kaedah pengekstrakan RNA yang telah diubahsuaikan khas untuk *S. binderi*, dapat menghasilkan RNA yang berkuantiti tinggi dengan pencemaran polisakarida yang minimum. mRNA yang cukup telah digunakan untuk membina perpustakaan cDNA utama dengan titer 9.2 x 10^5 pfu/ml. Sejumlah 2051 EST telah dijana dan dianalisis daripada perpustakaan cDNA yang telah diamplifikasi (titer 1.31 x 10^9 pfu/ml). EST dikenalpasti melalui perbandingan dengan pangkalan data peptida yang tidak berulang di NCBI. Sebanyak 82% EST telah digolongkan dalam kategori 'tidak diketahui' and 'kategori baru', yang berpotensi dan berkepentingan dalam penemuan gen-gen baru. Di samping itu, sebanyak 18% EST yang lain telah digolong dalam pelbagai kumpulan berdasarkan kepada fungsi putative EST tersebut, iaitu sintesis protin, tenaga, destinasi dan penyimpanan protin, metabolisma, struktur/pembahagian sel, penyakit/pertahanan, pemindahan isyarat, transkripsi dan serbaneka. Maklumat EST yang terkumpul dalam kajian ini, akan menyumbang ke arah pemahaman yang lebih mendalam mengenai biokimia dan biologi molecular *S. binderi* dan rumpai laut perang di masa akan datang.

ACKNOWLEDGEMENTS

Million thanks to Dr. Ho Chai Ling, my main supervisor, for giving me a good opportunity to pursue my studies in the field of molecular biology. Her knowledge, guidance, openness, enthusiasm, and inspiration will always serve to me as an example of the perfect supervisor. Not to forget Assoc. Prof. Dr. Raha Abdul Rahim and Prof. Dr. Phang Siew Moi, my co-supervisors, for their guidance and advice throughout my study.

I would like to extend my gratitude to my lab mates: Dr. Sudha, Yeen Yee, Seddon, Dang, Swee Sen and many others for their kind advice, assistance and support. Special thanks to Mr. Lee Weng Wah and Mr. Lee Yang Ping for assisting me in sequence analysis; members at the Algae Laboratory, University of Malaya, for their knowledge in seaweeds identification.

I would like to acknowledge the Malaysia Toray Science Foundation (MTSF) and the fundamental Research grant, the Ministry of Education, Malaysia, for providing financial support in this research project and Graduate Research Assistance fellowship.

Lastly, not to forget my family for their faith, love and emotional support. For that, I dedicate this dissertation to them.

I certify that an Examination Committee met on 20th July 2005 to conduct the final examination of Tony Wong Kok Min on his Master of Science thesis entitled "Optimization of RNA Extraction and Generation of Expressed Sequence Tags from *Sargassum binderi* (Sonder) J. Agardh" in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the candidate be awarded the relevant degree. Members of the Examination Committee are as follows:

Gulam Rusul Rahmat Ali, PhD
Professor
Faculty of Food Science and Technology
Universiti Putra Malaysia
(Chairman)

Maziah Mahmood, PhD
Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Internal Examiner)

Norihan Mohd Salleh, PhD
Associate Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Internal Examiner)

Rofina Yasmin Othman, PhD
Associate Professor
Institute of Biological Sciences
Universiti Malaya
(External Examiner)

GULAM RUSUL RAHMAT ALI, PhD
Professor/Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

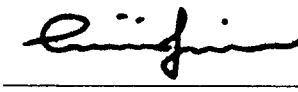
Date: 25 OCT 2005

This thesis submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science degree. The members of the Supervisory Committee are as follows:

Ho Chai Ling, PhD
Lecturer
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Chairman)

Raha Abdul Rahim, PhD
Associate Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Member)

Phang Siew Moi, PhD
Professor
Institute of Biological Sciences
Universiti Malaya
(Member)



AINI IDERIS, PhD
Professor/ Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 17 NOV 2005

DECRALATION

I hereby declare that the thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UPM or other institutions.

TONY WONG KOK MIN

Date: 19/10/2005

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

β	beta
λ	lambda
μg	microgram
μl	microliter
°C	degree centigrade
%	percentage
BLAST	Basic Local Alignment Search Tool
bp	base pairs
cDNA	complementary DNA
Cl	chloride
CsCl	cesium chloride
CTAB	hexacetyltrimethyl ammonium bromide
dATP	2'-deoxy-adenosine-5'-triphosphate
dCTP	2'-deoxy-cytidine-5'-triphosphate
DEPC	diethyl pyrocarbonate
dGTP	2'-deoxy-guanosine-5'-triphosphate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTPs	deoxynucleotides
DTT	dithiothreitol
dTTP	thymidine-5'-tryphosphate
EDTA	ethylenediaminetetraacetic acid
EtBr	ethidium bromide
g	gram

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	millimolar
mRNA	messenger RNA
NaCl	sodium chloride
NaOAc	sodium acetate
NaOH	sodium hydroxide
NCBI	National Center for Biotechnology Information
ng	nanogram
OD	optical density
PCR	polymerase chain reaction
pfu	plaque forming units
RNA	ribonucleic acid
RNase A	ribonuclease A
rpm	revolution per minute
SDS	sodium dodecyl sulphate
v/v	volume per volume
w/v	weight per volume

CHAPTER 1

INTRODUCTION

Seaweeds are macroalgae in the marine habitat. They can be used as food, sources for cosmetics, pharmaceuticals, soap, agar and textile. There are mainly three groups of seaweed: Phaeophyta (brown algae), Rhodophyta (red algae) and Chlorophyta (green algae) (Graham and Wilcox 2000), and each of them has its own important commercial values. The brown seaweeds are producer of alginate (Ertesvag and Valla 1998). Alginate is widely used in industry due to its viscosity and water binding capacity, and its ability to form thermostable gels with divalent cations (Ertesvag *et al.* 1998). *Sargassum binderi* is one of the most abundant brown seaweeds in Malaysia. Previous studies have shown that *S. binderi* contains high quality of alginate compared to commercially used seaweeds such as *Laminaria hyberborea*, *Macrocystis pyrifera* and *Ascophyllum nodosum*.

Therefore, it will be interesting to understand the biological functions of *S. binderi* through functional genomic studies. In this study, expressed sequence tag (EST) approach is chosen for the initial molecular studies on *S. binderi*, as very little genomic information is available. ESTs are partial sequences of cDNAs that can be used to characterize gene expression in organisms or tissues. These sequences or tags have been proven useful in many applications, such as recovery of full-length cDNA or genomic clones, discovery of novel genes, recognition of exons, delineation of protein families, development of genetic maps, identification of organism- or tissue-

specific genes, and investigation of unknown function (Lluisma and Ragan 1997). This study presents the first attempt in generating ESTs from *S. binderi*.

However, extraction of nucleic acids from seaweeds has always been problematic. During the nucleic acid extraction process, secondary metabolites and polysaccharides are always released after the disruption of cell (Kim *et al.* 1997). Besides, well established RNA extraction methods for higher plants may not work well in seaweeds, in which the seaweed polysaccharides may have different properties compared to higher plants. Therefore, optimization of RNA extraction from *S. binderi* is a necessary procedure in order to obtain high yield of RNA with minimum contamination.

The objectives of this study are to develop a RNA extraction method for *S. binderi* to obtain high yield of pure RNA, to construct a cDNA library from *S. binderi*, and to generate and analyze 1,000 ESTs from this seaweed for further functional genomic studies.

CHAPTER 2

LITERATURE REVIEW

2.1 Seaweed

Seaweeds are macroalgae or macroscopic members of the divisions Chlorophyta, Phaeophyta, and Rhodophyta living in the sea. They are plants visible to the naked eye, generally growing attached to solid substrate between and below the tide marks and remain stationary throughout life (Chapman 1979; Dawson 1956). Seaweeds rarely grow in the free floating state. However, there are sizable quantities of the brown alga *Sargassum* living in the free floating state in the Sargasso Sea northeast of the Caribbean and in the Gulf of Thailand (Dawson 1956).

The brown and red varieties are the more important seaweeds for commercial purposes. Seaweeds are phytosynthetic and able to manufacture sugar from water and carbon dioxide. Seaweeds contain chlorophyll, but this is clearly evident only in the green seaweeds. The colour may vary considerably according to environmental factors, such as the supply of nutrients, intensity of light, and time of the year. Red seaweeds may sometimes appear yellow, green or purple, and brown seaweeds may appear yellow, orange or greenish-black (Dawson 1956).

2.1.1 Classification

Although seaweeds are divided into red, brown and green seaweeds, they can not be differentiated by their colours. The classification is dependent on the presence of pigments, plastid organization and the properties of their cell wall. Table 2.1 shows the significant characteristics of seaweed in different divisions.

2.1.2 Morphology

Seaweeds are very plantlike in appearance, having root, stem, and leaf analogs in the form of anchoring holdfast, stipes, and blades (Linda and Lee 2000). A multicellular seaweed plant body is usually called thallus (Chapman 1979). A thallus may be a simple filament, a wide, broad sheet, a frond with flat or cylindrical divisions, or a clumped or close crust on rock, stone, shell or debris. Reproduction is by single cells produced over wide areas of the surface, in restricted patches, or on special branches (Alan 1977). The reproduction systems of seaweed may include vegetative propagation, asexual reproduction and sexual reproduction (Chapman 1979).

Table 2.1. Summary of the three macroalgae divisions and their significant characteristics (adapted from Harold and Michael 1978; David 2002)

Division	Phaeophyta	Rhodophyta	Chlorophyta
Common name	Brown algae	Red algae	Green algae
Pigments and plastid organization in photosynthetic species	Chlorophyll *a*, *c*; *β*-carotene, fucoxanthin and several other xanthophylls; 2-6 thylakoids/stack.	Chlorophyll *a*, (*d* in some florideophycidae); R- and C-phycocyanin, allophycocyanin; R- and B-phycoertthrin. *α*-, *β*-carotene, several xanthophylls; thylakoids single, not associated.	Chlorophyll *a*, *b*; *α*-, *β*-, and *γ*-carotenes, several xanthophylls; 2-5 thylakoids/stack.
Stored food	Laminaran (*β*-1,3-glucopyranoside, predominantly); mannitol.	Floridean starch (glycogen-like)	Starch (amylase and amylopectin) (oil in some).
Cell wall	Cellulose, alginic acid, and sulfated mucopolysaccharides (fucoidan).	Cellulose, xylans, several sulfated polysaccharides (galactans) calcification in some.	Cellulose in many (*β*-1,4-glucopyranoside), hydroxyproline glycosides; xylans and mannans; or wall absent; calcified in some.
Examples	*Fucus*, *Laminaria*, *Sargassum*	*Gracilaria*, *Porphyra*, *Chondrus*	*Enteromorpha*, *Ulva*, *Codium*

2.1.3 The importance of seaweed

Seaweeds have been harvested for many centuries for numerous uses. Since the time of Romans it has been used as a fertilizer, especially by farming communities living close to sea (Alan 1977).

At least 107 genera and 493 species of seaweeds have been recorded as being economically utilized worldwide (Tseng 1981). Three genera and three species of Chlorophyta, 10 genera and 13 species of Phaeophyta and four genera and six species of Rhodophyta are commercially cultivated in the Asia-Pacific region (Trono 1986). Seaweeds are mainly harvested for the phycocolloids (agar, carragreenan and alginic acid) (Phang 1984). The worldwide production of phycocolloids (first value is dry tonnage of raw material with the weight of phycocolloid in bracket): alginate 500,000 (18,000), carrageenan 200,000 (15,000), agar 180,000 (7,000) (Alan and Masao 1993). In Japan, the overall production of alginate is about 1,000-1,500 tons per year in 1994 (Subhuti 2002).

Agar is imported and marketed in Malaysia in four main forms: agar strips, bacteriological agar (powder), agar desserts (jellies) and flavoured powder mixes. Malaysia imported 172 tonnes of agar strips worth RM 6.55 millions in 1988, suggesting that a large domestic market exists to support the production and processing of seaweed in the country. However, competition from other countries in the region may be a serious constraint for the development of Malaysian seaweed processing and agar production (Jahara and Phang 1989).

2.2 Phaeophyta

The phaeophyta (brown algae) is the most complex forms found among the algae. There are a lot of differences compared to other algae (Table 1). The plant body of phaeophyta ranges from a millimeter in size or so to about 70 meters in length. They may be small, branched, attached, in filamentous forms, or have large plant bodies with certain portion similar to those found in higher plants. They are found most often firmly attached to various substrates, often with elaborate holdfast systems. In addition to these structures, which resemble roots, some forms have stem-like and leaf-like appendages. However, they are lack of vascular tissues of higher plants (Trainor 1978).

The colour of the brown algae (although colours may in fact vary from dark brown or golden brown to olive green), is due to an accessory carotenoid pigment, fucoxanthin, which masks the other pigments (Table 1) (Boney 1966). All the brown algae, with the exception of the Fucales, have an alternation of sporophyte and gametophyte generations (Dawson 1956). The cell wall matrix of marine algae typically consists of acidic polysaccharides. In the Phaeophyta, the major polysaccharide in the cell wall matrix is alginate (Hagen and Larsen 1997).

2.2.1 *Sargassum*

Genus *Sargassum* is under the division of Phaeophyta, order Fucales, and family Sargassaceae (Silva 1962). *Sargassum* is a large genus with more

than 150 species described, occurring in tropical, subtropical, and temperate zones of both hemispheres (Harold and Michael 1978). They are commonly called gulf weed and mostly are from tropical countries (Trainor 1978). *S. binderi* grow near to the seashore, especially in Malaysia, Philippines, India, and Australia (Misra 1966).

The *Sargassum* plant (Figure 2.1) is a flat, expanded sheet of cellular tissue (the thallus) that narrows towards the base, so forming the "stipe" or stalk, which is attached to the plant's habitat by the disc-shaped "holdfast". The thallus cells are in crowded rows with thick jelly-like cell walls, which give the plant flexibility and a degree of protection against the violent wave action. The plant is slippery, helping it to remain moist when uncovered and to prevent different specimens from chafing against and damaging each other when being swirled around by the incoming sea. The sliminess also helps the plant to maintain its elasticity, so that it can stretch and withstand the sucking and pulling of the sea movements and it is through exuding slime from its surface that it rids itself of calcium carbonate.

The thallus and its branches each have a conspicuous mid rib, on either side which are numerous oval-shaped air-filled bladders. Their purpose is to help buoy the plant when it is submerged, raising its extremities towards the water-surface and the light.