CHARACTERIZATION OF BACTERIAL COMMUNITIES ASSOCIATED WITH THE DINOFLAGELLATE, *Pyrodinium bahamense* var. *compressum*

SALLEY VENDA LAW

THESIS SUBMITTED IN FULLFILLMENT FOR DEGREE OF MASTER SCIENCE

BIOTECHNOLOGY RESEARCH INSTITUTE
UNIVERSITY MALAYSIA SABAH
2017

UNIVERSITI MALAYSIA SABAH

BORANG PENGESAHAN STATUS TESIS

JUDUL: **CHARACTERIZATION OF BACTERIAL COMMUNITIES ASSOCIATED WITH THE DINOFLAGELLATE, *Pyrodinium bahamense* var. *compressum***

IJAZAH : **MASTER OF SCIENCE (BIOTECHNOLOGY)**

Saya **SALLEY VENDA LAW**, sesi **2013-2014**, mengaku membenarkan tesis Sarjana ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat – syarat kegunaan seperti berikut:

1. Tesis in adalah hak milik Universiti Malaysia Sabah
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat Salinan untuk tujuan pangajian sahaja.
3. Perpustakaan dibenarkan membuat Salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (/):

[ ]	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA 1972)
[ ]	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
[/]	TIDAK TERHAD	

SALLEY VENDA LAW
MZ1321008T

Tarikh : 09 May 2017

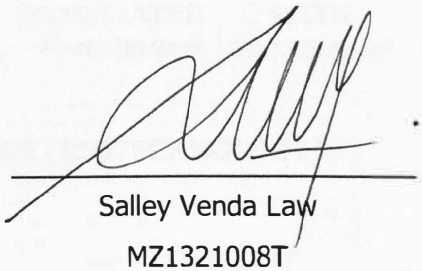
Disahkan Oleh,
NURULAIN BINTI ISMAIL
LIBRARIAN
UNIVERSITI MALAYSIA SABAH

(Tandatangan Pustakawan)

(Puan. Grace Joy Chin Wei Lie)
Penyelia

DECLARATION

I hereby declare that the material in this thesis is my own except for quotations, excerpts, equations, summaries, and references, which have been duly acknowledged.



15 March 2017

Salley Venda Law

MZ1321008T

CERTIFICATION

NAME	:	**SALLEY VENDA LAW**
MATRIC NO	:	**MZ1321008T**
TITLE	:	**CHARACTERIZATION OF BACTERIAL COMMUNITIES ASSOCIATED WITH THE DINOFLAGELLATE, *Pyrodinium bahamense* var. *compressum***
DEGREE	:	**MASTER OF SCIENCE (BIOTECHNOLOGY)**
DATE OF VIVA	:	**15 MARCH 2017**

CERTIFIED BY

Signature

1. **SUPERVISOR**
   Madam Grace Joy Chin Wei Lie _____

2. **CO-SUPERVISOR**
   Assoc. Prof. Dr. Kenneth Francis Rodrigues _____

3. **CO-SUPERVISOR**
   Prof. Datin Dr. Ann Anton _____

ACKNOWLEDGEMENTS

I owe my gratitude to all those people who made this thesis possible and because of whom my master degree experience has been one that I will cherish forever. First and foremost, I would like to thank God for providing me the opportunity and granting me the capability to compete this research work.

I would like to express my deepest gratitude and special appreciation to my supervisory committee, Prof. Datin Dr. Ann Anton, Assoc. Prof. Dr. Kenneth Francis Rodrigues, and Madam Grace Joy Chin Wei Lie, for their immense knowledge, thoughtful guidance and support at different stages of my research work that lead to the completion of this thesis. I would like to acknowledge all the Biotechnology Research Institute lecturers for allowing me to access to the laboratory and research facilities.

Besides, I would also like thank all members of UHABs, especially Mr. Yusdi, Ms. Audrey Rose and Ms. Adriane, for their assistance and support during the whole period of the study especially for their help in *Pyrodinium bahamense* var. *compressum* culture laboratory establishment. Without they precious support it would not be possible to conduct this research. My sincere thanks also go to the bioinformatics services provider from Bioeasy Company, especially Jaeyres Jani who was involved in the bioinformatics analysis project. Thank you for the countless guidance and kind assistance from the beginning process of analysing data until the finalized data result of this research project.

Most importantly, none of this would have been possible without the love and patience of my family. Thank you for the unfailing moral support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. Last but not least, I would like to thank to my fellow friends who have helped me in many different way directly or indirectly. This accomplishment would not have been possible without them. Thank you very much.

SALLEY VENDA LAW
15 March 2017

ABSTRACT

Toxic dinoflagellate, *Pyrodinium bahamense* var. *compressum* (PBVC) is the causative agent of paralytic shellfish poisoning (PSP) in Sabah harmful algal bloom (HAB) occurrence. Pure unialgal cultured strain CC-UHABS-040(M) harvested from the late exponential phase used for bacteria isolation. Study aims to characterize bacterial communities associated with PBVC by using two molecular approaches; the 16S rDNA culture-dependent and culture-independent 16S rRNA metagenomics analysis. The potential putative PST (Paralytic shellfish toxin)-producing bacterium, PBVC088 selected for whole genome sequencing analysis to produce draft genome of *Ruegeria* sp. and further identify the *sxt* (saxitoxin) genes in their genome. Colony morphology analysis revealed the 74 bacterial colonies were small punctiform or large circular, smooth-textured and non-pigmented. The culture-dependent analysis revealed bacteria diversity was limited to gamma-proteobacteria and alpha-proteobacteria, where it is predominated by *Alteromonas* sp. and *Ruegeria* sp. The taxonomic profile analysis using culture-independent approach showed high bacterial diversity in Sabah PBVC culture, which was classified into 20 classes, 43 orders, 60 families, and 105 genera. In addition, bacterium *Ruegeria* sp. (PBVC088) exhibited a draft genome size of 5, 784, 660 bp with a G+C content of 65 %, and containing 5, 640 protein coding sequences (CDSs). BLASTP sequence similarity search of saxitoxin proteins (expectation value, E-value < $1e^{-5}$) against the 26 putative *sxt* genes of the toxic cyanobacterium, *Cylindrospermopsis raciborskii* T3 successfully identified eleven putative *sxt* candidate genes. The eleven *sxt* genes (*sxtA*, *sxtB, sxtF/M, sxtH/T, sxtS, sxtU, sxtV, sxtW*, and *sxtZ*) found to contain similar conserved domains as in the cyanobacterial domain, *C. raciborskii* T3. The finding suggesting, genes (enzymes) associated with the STX biosynthesis pathway exist in the bacterial genome PBVC088, which most likely involved either in the biosynthesis of the final compound or as the precursor of the biosynthesis pathway. High-throughput genome analysis of the associated bacterial communities helps elucidate the hypothesis that the associated bacteria may or may not be involve in the PST production and subsequently help reveal their potential function.

ABSTRAK

PENCIRIAN KOMUNITI – KOMUNITI BAKTERIA BERKAITAN DENGAN DINOFLAGELAT, Pyrodinium bahamense var. *compressum*

Dinoflagelat toksik, Pyrodinium bahamense var. compressum (PBVC) *adalah agen penyebab keracunan paralitik kerang* (PSP) *dalam kejadian ledakan alga berbahaya di* Sabah*. Kultur unialga tulen strain* CC-UHABS-040(M) *dituai pada fasa eksponen akhir digunakan untuk pengasingan bakteria. Kajian bertujuan untuk mencirikan komuniti – komuniti bakteria yang berkaitan dengan* PBVC *dengan mengunakan dua pedekatan molekular;* 16S rDNA *bersandar-kultur and tak bersandar-kultur* 16S rRNA *metagenomik analisis. Potensi putatif bakteria pengeluar* PST *(toksin paralitik kerang),* PBVC088 *dipilih untuk analisis penjujukan keseluruhan genom bagi menghasilkan draf genom* Ruegeria sp. *dan seterusnya mengenal pasti gen – gen* sxt (saxitoxin) *dalam genom mereka. Analisi coloni morfologi mendedahkan 74 coloni bakteria adalah punktiform kecil atau bulat besar, licin bertekstur, bukan-berpigmen. Analisis bersandar-kultur mendedahkan kepelbagaian bakteria adalah terhad kepada gamma-proteobakteria dan alfa-proteobakteria dengan didominasi oleh* Alteromonas sp*. dan* Ruegeria sp*. Analisis profil taksonomi menggunakan pendekatan tak bersandar-kultur menujukkan kepelbagian bakteria yang tinggi dalam kultur* Sabah PBVC *yang dikelaskan kepada* 20 kelas, 43 ordo, 60 famili, *dan* 105 genera*. Bakteria* Ruegeria sp. *mempamerkan saiz draf genom* 5, 784, 660 bp *dengan kandungan* G+C 65 %, *dan mengandungi* 5, 640 *jujukan-jujukan protein pengekodan (CDSs). BLASTP jujukan carian persamaan protein saxitoxin (nilai jangkaan, E-nilai < $1e^{-5}$) terhadap 26 putative gen saxitoksin (sxt) daripada toksik cyanobakteria,* Cylindrospermopsis raciborskii T3*, berjaya mengidentifikasi sebelas putatif calon gen sxt. Sebelas sxt gen – gen (sxtA, sxtB, sxtF/M, sxtH/T, sxtS, sxtU, sxtV, sxtW, dan sxtZ) itu didapati mengandungi domain abadi sama seperti dalam domain cyanobakteria* C. raciborskii T3*. Penemuan kajian mencadangkan gen-gen (enzim-enzim) yang dikaitkan dengan laluan* STX *biosintesis wujud dalam genom bakteria,*PBVC088*, yang berkemungkinan besar terlibat sama ada dalam biosintesis kompaun terakhir atau sebagai prekursor. Kendalian genom analisis yang tinggi terhadap komuniti* PBVC *membantu menrungkaikan hipothesis bahawa bakteria yang dikaitkan mungkin atau mungkin tidak terlibat dalam pengeluaran* PST *dan seterusnya membantu mendedahkan potensi fungsi mereka.*

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

α	-	alpha
B	-	beta
D	-	delta
ε	-	epsilon
γ	-	gamma
%	-	Percent
°C	-	Degree Celsius
µL	-	microlitre
µm	-	micrometer
µM	-	micromolar
ATM	-	Amplicon Tagment Mix
BLASTN	-	Basic local alignment search tool for nucleotide
BLASTP	-	Basic local alignment search tool for protein
Bp	-	Base pair
CDD	-	Conserved Domain Database
DMSP	-	Dimethylsulfoniopropionate
dH_2O	-	Distilled water
E-value	-	Expectation value
EBI	-	European Bioinformatics institute
EDTA	-	Ethylenediamine tetraacetic acid
ELISA	-	Enzyme-linked immunosorbent assay
EtBr	-	Ethidium bromide
EtOH	-	Ethanol
Faa	-	Protein FASTA file of the translated CDS sequences.
Ffn	-	Nucleotide FASTA file of all the annotated sequences
Fna	-	Nucleotide FASTA file of the Assembled sequences
G	-	Gram
gDNA	-	genomic deoxyribonucleic acid
g/mol	-	gram over molar
GTX	-	gonyautoxin
HAB	-	Harmful algal bloom
HGT	-	Horizontal Gene Transfer
HPLC	-	High performance liquid chromatography
Hr	-	Hour
HT1	-	Hybridization buffer
I	-	Iodine
IDBA-UD	-	Iterative De Bruijn Graph *De Novo* Assembler
Kb	-	kilobase pair
M	-	Molar
MEGAN	-	Metagenome analyzer
Mg	-	Miligram
MgCl2	-	Magnesium

Min	-	Minutes
mL	-	millilitre
Mm	-	Millimeter
mM	-	millimolar
Nm	-	Nanometer
Ng	-	Nanogram
nM	-	nanomolar
NaCl	-	Sodium chloride
NT	-	Neutralize Tagment Buffer
NPM	-	Nextera PCR Master Mix
NGS	-	Next generation sequencing
NaOH	-	Sodium hydroxide
NCBI	-	National center for biotechnology information
neoSTX	-	neosaxitoxin
NO_3	-	Nitrate nitrogen
OTUs	-	Operation Taxonomic Units
Q	-	Quality score
qPCR	-	quantitative Polymerase chain reaction
P	-	Probability error
PCR	-	Polymerase chain reaction
PRODIGAL	-	Prokaryotic Dynamic Programming Genefinding Algorithm
PBVC	-	*Pyrodinium bahamense* var. *compressum*
PEAR	-	Paired-end read merger
pM	-	pico molar
PO_4	-	Phosphate
PSP	-	Paralytic Shellfish Poisoning
PST	-	Paralytic Shellfish Toxin
Rpm	-	revolution per minute
rRNA	-	ribosomal ribonucleic acid
rDNA	-	ribosomal deoxyribonucleic acid
S	-	Second
SRA	-	Sequence Read Archive
SSU	-	Small subunit
STX	-	Saxitoxin
Sxt	-	Saxitoxin
TD	-	Tagment DNA Buffer
TBE	-	Tris Borate EDTA
TE	-	Tris-HCl EDTA
Tris-HCl	-	Tris (hydroxymethyl) aminomethane hydrochloride
U	-	Unit
UV	-	Ultra violet
V	-	Variable region
v/v	-	Volume over volume
w/v	-	Weight over volume
WGS	-	Whole genome sequencing

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1. Research Background

Harmful algal bloom (HAB) or also known as "red tide" is a common phenomenon in coastal waters worldwide. Under several factors such as nutrient concentrations, weather conditions, change in seawater parameters and geomorphology location (Tilstone *et al.,* 1994; Tan *et al.,* 2006; Adam *et al.*, 2011), a unicellular harmful alga may proliferate and/or aggregate to form dense concentrations of cells or "blooms" resulting in a massive HAB outbreak (Van Dolah, 2000). The outbreak can have severe negative impacts on the environment and human health through consumption of contaminated seafood products, and may cause substantial economic losses to the aquaculture, fisheries and tourism industries (Van Dolah, 2000; Gedaria *et al.,* 2007). Of the estimated 60 – 80 HAB species that reported as toxic or harmful, only 10 – 12 dinoflagellates species primarily responsible for the current HABs outbreak worldwide (Hallegraeff, 1993; Smayda, 1997).

The production of saxitoxins (STXs) found in several dinoflagellates species from the genus *Alexandrium* spp., *Gymnodinium catenatum* and *Pyrodinium bahamense* var. *compressum* (PBVC) (Oshima *et al.,* 1993; Usup *et al.,* 1994; Doucette and Trick, 1995). The thecate, chain-forming dinoflagellate PBVC is the main causative organism responsible for paralytic shellfish poisoning (PSP) in Southeast Asian countries such as Malaysia and the Philippines as well as the Pacific coastline of Central America (Orellana-Cepeda *et al.,* 1998; Usup *et al.,* 2002). Harmful algal blooms have regularly occurred over the past three decades in the coastal waters of western Sabah with the first reported case of PBVC blooms in 1976 (Roy, 1977; Anton *et al.,* 2008). Relatively more severe HAB outbreaks have been detected in recent years along the west coast of Sabah, which indirectly caused economic losses to the aquaculture industries in Sabah as well as resulted in human illnesses and fatalities (Usup *et al.,* 2012). Globally, the distribution and frequency of HAB events have been increasing with 2,000 PSP cases reported per year at a human mortality rate of 15 % (Hallegraeff, 1995).

Paralytic shellfish poisoning (PSP) is a potential fatal neurological disorder caused by paralytic shellfish toxins (PSTs) or a group of neurotoxins collectively known as saxitoxins (STXs). These STXs act by blocking movement of sodium ions through nerve cell membranes that subsequently stop the flow of nerve impulses and thus cause symptoms of PSP (Mosher *et al.,* 1964). Muscular paralysis would ensue and potentially death from respiratory failure in chronic poisoning cases (Tan and Ransangan, 2015). In Sabah, the most recent severe PSP cases were reported in January 2013, with three casualties and 43 hospitalizations (The Star, 2013).

In both marine and freshwater environments, symbioses between bacteria and algae are commonly observed, and some studies hypothesized that the interaction potentially is responsible for toxin production during a HAB event (Gallacher and smith, 1999; Alavi *et al.,* 2001; Alverca *et al.,* 2002; Córdova *et al.,* 2003; Azanza *et al.,* 2006). Additionally, bacteria-algae dynamics during HABs have been postulated to play an important role in regulating the processes of algal bloom initiation, maintenance, and decline (Doucette, 1995; Ferrier *et al.,* 2002). Tobe (2003) suggested that effects of the bacteria associated with toxic dinoflagellates could be either direct or indirect in toxin production. Despite the growing volume of literatures on the capabilities of these bacteria to metabolize, produce, and modify toxins autonomously, it has yet to be confirmed that these toxin-producing bacteria are capable of autonomous PSP toxin synthesis (Córdova *et al.,* 2003).

Several hypotheses on the interaction between algae and bacteria have been suggested, these including nutrient exchange such as dimethylsulfoniopropionate (DMSP) metabolism (Buchan *et la.,* 2014), signal transduction and horizontal gene transfer (Kellmann *et al.,* 2008b). Recently, three theories have been proposed including a polyphyletic origin of the involvement of symbiotic bacteria in dinoflagellates, convergent evolution of analogous STX products or pathways in both symbiotic bacteria and host dinoflagellate, and spreading through horizontal gene transfer (HGT) (Orr *et al.,* 2013a). Despite the increasing reports on the tight association, but the phylogenetic origin of genes involved in the biosynthesis of STX remained elusive.

Previous studies reported that the diversity of bacterial communities associated with toxic dinoflagellates commonly dominated by two bacterial phyla, the Proteobacteria and the Bacteriodetes (Alverca *et al.,* 2002; Azanza *et al.*, 2006; Chin *et al.,* 2013). An earlier study on the identification of bacteria associated with Sabah PBVC uncovered a total of 16 bacterial isolates whose identities limited to the phyla Proteobacteria and Actinobacteria obtained from a clonal culture of PBVC (Chin *et al.,* 2013). Another study carried out by Azanza *et al.* (2006) on bacterial diversity from Philippines PBVC have identified PSTs-secreting endosymbionts bacteria including *Moraxella* spp., *Erythrobacter* spp., and *Bacillus* spp. (Azanza *et al.,* 2006). Several other studies had also successfully identified other PST-producing bacteria such as *Moraxella* sp., *Alteromonas* sp. and *Roseobacter* sp. from other toxic dinoflagellates species *Alexandrium* spp. and *G. catenatum* (Kodama and Ogata, 1988; Gallacher and Smith, 1999; Córdova *et al.,* 2002). Hence, these bacteria continue to be only putatively toxic (Groben *et al.,* 2000).

Studies carried out to identify symbionts of PBVC have hitherto been limited to culture-dependent methods that are vulnerable to oversights of scarce or uncultivable taxa (Azanza *et al.,* 2006; Chin *et al.,* 2013). Therefore, the experimental design of this study would enable coverage of these uncultivable members of the bacterial communities via high-throughput 16S metagenomics sequencing technology. Recently, next generation sequencing technologies have advanced rapidly and Illumina Miseq sequencing, in particular, has become a popular platform since it can generate millions of sequence reads of partial 16S rRNA genes that is able to accommodate the throughput demands of ecological studies on environmental microbiota at a relatively lower cost (Kennedy *et al.,* 2010; Chaudhary *et al.,* 2015).

This study aimed to determine the diversity of bacteria associated with the toxic dinoflagellate, PBVC for both culturable and non-culturable bacteria using traditional and molecular taxonomic techniques. Estimation of bacterial diversity based on colony morphology and 16S rDNA characterization were applied as a preliminary screening tool, following which next generation sequencing platform for metagenomics sequencing was employed to analyze the composition of the microbial communities. In addition to this, whole genome sequencing of a putative PST-producing bacterium associated with PBVC was conducted to determine if the selected bacterium possesses

the repertoire of genes responsible for the biosynthesis of the toxins. Information from this study would facilitate the understanding of the bacterial diversity in the toxic blooms and as a means to assess the role of the associated bacteria in toxin production.

1.2. Objectives Of the Research

This study embarked upon the following objectives:

1) To characterize the bacterial diversity associated with PBVC based on colony morphology and culture-dependent 16S rRNA gene.
2) To determine the taxonomic profile of the bacterial diversity associated with PBVC based on culture-independent 16S metagenomics sequencing.
3) To identify the genes involved in the biosynthesis of toxins from a putative PST-producing bacterium, PBVC088 by sequencing the whole genome of the associated bacterium.

1.3. Significance Of Study

Identification and characterization of bacterial taxa associated with PBVC isolate from Sabah further add to current knowledge on their biodiversity which would also include clades that are previously obscured due to limitations of techniques used in previous inquiries. Thus, this study would also demonstrate the feasibility of engaging massively parallel sequencing approach to comprehensively study complex microbial communities of Sabah PBVC. On the other hand, decoding the whole genome of a putative PST-producing bacterium will provide a window of opportunity to understand the role of the associated bacteria in toxin production during HABs. In the long run, the findings generated would contribute to the larger efforts of implication in understanding the STX biosynthesis pathway and may have further potential application in medical applications.