

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

**SALLEY VENDA LAW**



**UMS**

**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 pengajian 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)

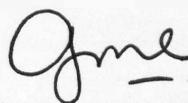
TIDAK  
TERHAD

  
**SALLEY VENDA LAW**  
**MZ1321008T**

Disahkan Oleh,  
**NURULAIN BINTI ISMAIL**  
LIBRARIAN  
UNIVERSITI MALAYSIA SABAH

  
(Tandatangan Pustakawan)

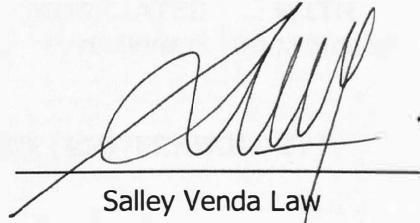
Tarikh : 09 May 2017

  
(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



UMS  
UNIVERSITI MALAYSIA SABAH

## 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

UMMS  
UNIVERSITI MALAYSIA SABAH

Signature

**1. SUPERVISOR**

Madam Grace Joy Chin Wei Lie

Handwritten signature of Madam Grace Joy Chin Wei Lie, written in black ink on a white background.

**2. CO-SUPERVISOR**

Assoc. Prof. Dr. Kenneth Francis Rodrigues

Handwritten signature of Assoc. Prof. Dr. Kenneth Francis Rodrigues, written in black ink on a white background.

**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 *Pyrodictum bahamense* var. *compressum* culture laboratory establishment. Without their 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 menggunakan dua pendekatan 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 menunjukkan kepelbagaian bakteria yang tinggi dalam kultur Sabah PBVC yang dikelaskan kepada 20 kelas, 43 ordo, 60 famili, dan 105 genera. Bacteria *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 saxitoxin (*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 menrunjkaikan hipotesis bahawa bakteria yang dikaitkan mungkin atau mungkin tidak terlibat dalam pengeluaran PST dan seterusnya membantu mendedahkan potensi fungsi mereka.

## TABLE OF CONTENT

TITLE	Page
DECLARATION	i
CERTIFICATION	ii
ACKNOWLEDGEMENTS	iii
ABSTRACT	iv
<i>ABSTRAK</i>	v
TABLE OF CONTENT	vi
LIST OF TABLES	vii
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xiii
LIST OF APPENDICES	xvii
<b>CHAPTER 1: INTRODUCTION</b>	
1.1 Research Background	1
1.2 Objectives of the Research	4
1.3 Significance of Study	5
<b>CHAPTER 2: LITERATURE REVIEW</b>	
2.1 The Dinoflagellate, <i>Pyrodinium bahamense</i> var. <i>compressum</i>	6
2.2 Overview of Harmful Algal Blooms (HABs)	8
2.3 HABs in Sabah, Malaysia	9
2.4 Paralytic Shellfish Poisoning (PSP) and Paralytic Shellfish Toxins (PSTs)	12
2.4.1 Overview of PSP	12
2.4.2 Chemical Characteristic of PSTs	14
2.4.3 Detection of PSTs	17

2.5	Toxic Dinoflagellate and Its Associated / Symbiotic Bacteria	20
2.5.1	Bacterial Association with Dinoflagellates	20
2.5.2	Toxins Associated with Bacterial Symbionts of Dinoflagellates	25
2.5.3	Molecular Characterization of Bacteria Associated with HABs	28
2.6	Saxitoxin (STX) Pathway	29
2.6.1	Biosynthetic Pathways Associated with STX Production	29
2.6.2	Dinoflagellate Genes Associated with STX Biosynthesis	37
2.6.3	Bacterial Genes Associated with STX Biosynthesis	41
2.7	16S rDNA Metagenomics Analysis of Bacteria Associated with Dinoflagellates	44
2.8	Whole Genome Sequencing	48

### **CHAPTER 3: METHODOLOGY**

3.1	Study Area and Sample Collection	54
3.2	Establishment of Strain CC-UHABS-040(M) <i>in vitro</i>	54
3.3	Estimating the Growth Characteristics of Strain CC-UHABS-040(M)	54
3.4	Determination of the toxicity of strain CC-UHABS-040(M)	55
3.5	Isolation of Bacteria Associated with Strain CC-UHABS-040(M)	56
3.6	Gram-Staining	57
3.7	Ribotyping of bacteria Associated with Strain CC-UHABS-040(M)	57
3.7.1	Extraction of Genomic DNA	58
3.7.2	Polymerase Chain Reaction (PCR) Amplification of the 16S ribosomal DNA (rDNA)	59
3.7.3	16S rDNA Sequence Determination of Bacteria Associated with PBVC	60
3.8	16S Metagenomic Sequencing of Bacteria Associated with PBVC	60
3.8.1	Extraction of Microbial DNA Associated with PBVC	61

3.8.2	Amplicon Primers preparation	62
3.8.3	16S Metagenomic Sequencing	63
3.8.4	Data Analysis of Metagenomic Sequencing	69
3.9	Determination of <i>sxt</i> genes Involved in STX-biosynthesis in the Putatively PST-Producing <i>Ruegeria</i> sp. based on Whole Genome Sequencing	70
3.9.1	Sequencing of the <i>Ruegeria</i> sp. Genome	70
3.9.2	Library Quantification, Normalisation and Pooling	73
3.9.3	Denaturation of DNA Library	74
3.9.4	Combining Sample Library and Phix Control	74
3.9.5	Sample Loading	74
3.9.6	Next Generation Sequencing Data Analysis	75
<b>CHAPTER 4: RESULT AND DISCUSSION</b>		
4.1	Algal Culturing	77
4.2	Growth Rate Estimation	79
4.3	Determination of Toxicity	81
4.4	Isolation and Characterization of Culturable Bacteria Associated with the Cultured PBVC	84
4.5	Gram-Staining of the Bacterial Isolates	85
4.6	Identification of Bacterial Isolates Using Partial 16S rDNA Sequence	86
4.6.1	Isolation of Genomic DNA	86
4.6.2	PCR Amplification and Capillary Sequencing of the 16S rDNA	87
4.6.3	Taxonomic Assignment	88
4.7	Biodiversity Assessment of Bacterial Communities Associated with Sabah PBVC through Metagenomics Approach	94
4.7.1	Comparative Richness and Relative Abundance based on Partial 16S rDNA Metagenomics Analysis	96

4.7.2	Associated Bacteria in Sabah PBVC Culture	103
4.7.3	Limitation of the Experimental Approach	103
4.8	Whole-Genome Sequencing of a Putative PST-Producing Toxin, <i>Ruegeria</i> sp. (PBVC088)	104
4.8.1	Identification of saxitoxin proteins	109
<b>CHAPTER 5:</b>	<b>CONCLUSION AND FUTURE STUDIES</b>	120
<b>REFERENCES</b>		124
<b>APPENDICES</b>		144



**UMS**  
UNIVERSITI MALAYSIA SABAH

## LIST OF TABLES

	Pages
Table 2.1: Chemical formulae of the functional groups (R1 – R4) associated with several types of paralytic shellfish toxins	17
Table 2.2: Timeline of evidences supporting the hypothesis on bacterial origin of STX production reported from 1979 -2015	27
Table 2.3: The <i>sxt</i> genes of <i>C. raciborskii</i> T3 and their putative functions and catalytic product	32
Table 2.4: Quality score and error probabilities based on calculation from formula above	49
Table 3.1: Full length primer sequences	62
Table 3.2: Illumina overhang adapter nucleotide sequences	62
Table 3.3: Setup of PCR reaction to amplify the first amplicon	64
Table 3.4: PCR reaction used per reaction to generate TruSeq Index tubes for metagenome sample	66
Table 3.5: qPCR reaction mix for DNA library and six DNA standards	67
Table 3.6: PCR amplification reaction used per reaction for bacterial genome sequencing	72
Table 3.7: Annotated files description	76
Table 4.1: BLAST similarity results of rDNA sequences of the 74 isolates against NCBI database showing the closest relative (domain; phylum; class; order; family; genus; species), the accession number of the reference sequence in Genbank, and the range of percentage similarities.	89
Table 4.2: The ten phyla present in the 16S metagenomics analysis and the sources with references from previous studies.	100
Table 4.3: Major bacteria genera detected in the bacteria communities of Sabah PBVC and their sources with references from previous studies.	101
Table 4.4: Vital statistics of the <i>de novo</i> genome assembly output for <i>Ruegeria</i> sp. bacterium PBVC088	105

Table 4.5:	BLASTP sequence similarity search of genome against the 26 putative <i>sxt</i> genes of STX-producing cyanobacteria, <i>C. raciborskii</i> T3 and their accession numbers	110
Table 4.6:	Domain structure of the <i>sxt</i> -related proteins in PBVC088 compared to <i>Cylindrospermopsis raciborskii</i> T3 cyanobacterial <i>sxt</i> protein domain.	111
Table 4.7:	Conserved domains identified in putative candidate genes by BLAST search and CD-Search web service	113



UMS  
UNIVERSITI MALAYSIA SABAH

## LIST OF FIGURES

	Pages
Figure 2.1: Two fused cells PBVC forming a chain, observed under light microscope at 40 X magnification.	8
Figure 2.2: Map showing six locations (highlighted in red colour with numbers 1 – 6) with notably frequent bloom occurrences along the coastal waters of Sabah	10
Figure 2.3: The proposed transmembrane arrangement of the $\alpha$ -subunit of Na <sup>+</sup> channels. The pore is represented in red, the voltage sensors in yellow and the inactivation gate in blue. PSP is mediated by the interaction and blockage of Site 1 by STX.	15
Figure 2.4: Chemical structure of an STX (refer Table 2.1 for functional groups)	16
Figure 2.5: A summary of the metabolic features of Roseobacter bacteria that are involved in their interactions with host phytoplankton	24
Figure 2.6: Structural organization of the 35 kb <i>sxt</i> gene cluster from <i>C. raciborskii</i> T3 with its respective encode enzymes	30
Figure 2.7: The revised pathway for STX biosynthesis and the putative enzymes involved in the cyanobacteria species of STX-producing cyanobacteria, <i>C. raciborskii</i> T3	35
Figure 2.8: The three hypothesized theories of STX evolution in dinoflagellates. The dinoflagellate and cyanobacterial lineages are highlighted in red rectangles.	42
Figure 2.9: Bacterial 16S rRNA gene containing 9 hypervariable region (HVR) (V1-V9)	45
Figure 3.1: General workflow of NGS-based 16S rDNA sequencing	60
Figure 3.2: Diagram illustrating the summarised workflow for 16S Metagenomics sequencing	63
Figure 3.3: Diagram illustrates the workflow of bacterial DNA Whole genome sequencing using the Nextera XT kit	71
Figure 4.1: Pure unialgal culture of Sabah PBVC in sterile 500 ml cotton-plugged conical flask enriched with seawater-based f/2 medium.	77

Figure 4.2:	Versatile environmental test chamber used for maintaining the PBVC culture	78
Figure 4.3:	Growth curve of PBVC cultured in f/2 medium for 30 days in which the early exponential phase was from day 6 to 10, late exponential from day 12 to 16, and stationary or death phase occurring between day 21 to 30.	80
Figure 4.4:	HPLC Chromatogram profile of the Standard STX toxin (A <sub>1</sub> ) and GTX1-6 (B <sub>1</sub> ) and the cultured PBVC collected during the late exponential phase (A <sub>2</sub> ) for STX detection and (B <sub>2</sub> ) for GTX1-6 detection.	82
Figure 4.5:	A marine agar plate with 50 – 100 bacterial colonies associated with cultured PBVC from the late exponential phase (Day 13) plated after 10 <sup>-3</sup> serial dilution.	84
Figure 4.6:	The bacterial isolates with spiral or spore-forming and rod-shaped bacteria, Figure 4.6 (A) is the isolate PBVC096 viewed under 40 X magnification and Figure 4.6 (B) is the isolate PBVC097 viewed under 100 X magnification.	86
Figure 4.7:	Genomic DNA isolated from PBVC bacterial isolates viewed using 1.0% (w/v) agarose gel stained with ethidium bromide. Lane M: Lambda <i>HindIII</i> DNA ladder; Lanes 1 – 10: PBVC061; PBVC062; PBVC063; PBVC064; PBVC065; PBVC067; PBVC068; PBVC069; PBVC058; PBVC059.	87
Figure 4.8:	PCR amplification of 16S rDNA (~464 bp) for nine bacterial isolates. Lane M: 100 bp DNA Ladder; Lanes 11 – 19: PBVC026; PBVC041; PBVC051; PBVC055; PBVC059; PBVC067; PBVC072; PBVC076; PBVC084; Lane C: negative control.	88
Figure 4.9:	Genus-level distribution of the bacterial isolates	90
Figure 4.10:	The diversity curve of the sample investigated in this study	95
Figure 4.11:	The frequency (in log scale) of the number of reads assigned at the Family level	95
Figure 4.12:	Taxonomic affiliation of 16S rRNA metagenomic sequences. The phylogenetic distribution for (A) the overall population and (B) Proteobacteria	97
Figure 4.13:	Output summary of functional distribution of protein-coding genes in the PBVC088 from RAST server	106

- Figure 4.14: SNP Phylogenies tree was constructed using parsimony algorithm based core SNP matrix inferred from the closed genome of Roseobacter clade from NCBI database and the draft genome of *Ruegeria* PBVC088 (highlighted in red colour) by kSNP analysis. Bootstrap values (100 replicates) are reported above each node. 108
- Figure 4.15: The revised pathway for STX biosynthesis by Kellmann *et al.*, (2008a) and the highlighted in red colour showed the potential steps (step number 1,2,4,6,7,8 and 10) of the involvement of bacterial genes (PBVC088) in the STX-biosynthesis pathway. 115



UMS  
UNIVERSITI MALAYSIA SABAH

## LIST OF ABBREVIATIONS

<b>α</b>	- alpha
<b>β</b>	- beta
<b>Δ</b>	- delta
<b>ε</b>	- epsilon
<b>γ</b>	- gamma
<b>%</b>	- Percent
<b>°C</b>	- Degree Celsius
<b>μL</b>	- microlitre
<b>μm</b>	- micrometer
<b>μM</b>	- micromolar
<b>ATM</b>	- Amplicon Tagment Mix
<b>BLASTN</b>	- Basic local alignment search tool for nucleotide
<b>BLASTP</b>	- Basic local alignment search tool for protein
<b>Bp</b>	- Base pair
<b>CDD</b>	- Conserved Domain Database
<b>DMSP</b>	- Dimethylsulfonypropionate
<b>dH<sub>2</sub>O</b>	- Distilled water
<b>E-value</b>	- Expectation value
<b>EBI</b>	- European Bioinformatics institute
<b>EDTA</b>	- Ethylenediamine tetraacetic acid
<b>ELISA</b>	- Enzyme-linked immunosorbent assay
<b>EtBr</b>	- Ethidium bromide
<b>EtOH</b>	- Ethanol
<b><i>Faa</i></b>	- Protein FASTA file of the translated CDS sequences.
<b><i>Ffn</i></b>	- Nucleotide FASTA file of all the annotated sequences
<b><i>Fna</i></b>	- Nucleotide FASTA file of the Assembled sequences
<b>G</b>	- Gram
<b>gDNA</b>	- genomic deoxyribonucleic acid
<b>g/mol</b>	- gram over molar
<b>GTX</b>	- gonyautoxin
<b>HAB</b>	- Harmful algal bloom
<b>HGT</b>	- Horizontal Gene Transfer
<b>HPLC</b>	- High performance liquid chromatography
<b>Hr</b>	- Hour
<b>HT1</b>	- Hybridization buffer
<b>I</b>	- Iodine
<b>IDBA-UD</b>	- Iterative De Bruijn Graph <i>De Novo</i> Assembler
<b>Kb</b>	- kilobase pair
<b>M</b>	- Molar
<b>MEGAN</b>	- Metagenome analyzer
<b>Mg</b>	- Miligram
<b>MgCl<sub>2</sub></b>	- Magnesium

<b>Min</b>	- Minutes
<b>mL</b>	- millilitre
<b>Mm</b>	- Millimeter
<b>mM</b>	- millimolar
<b>Nm</b>	- Nanometer
<b>Ng</b>	- Nanogram
<b>nM</b>	- nanomolar
<b>NaCl</b>	- Sodium chloride
<b>NT</b>	- Neutralize Tagment Buffer
<b>NPM</b>	- Nextera PCR Master Mix
<b>NGS</b>	- Next generation sequencing
<b>NaOH</b>	- Sodium hydroxide
<b>NCBI</b>	- National center for biotechnology information
<b>neoSTX</b>	- neosaxitoxin
<b>NO<sub>3</sub></b>	- Nitrate nitrogen
<b>OTUs</b>	- Operation Taxonomic Units
<b>Q</b>	- Quality score
<b>qPCR</b>	- quantitative Polymerase chain reaction
<b>P</b>	- Probability error
<b>PCR</b>	- Polymerase chain reaction
<b>PRODIGAL</b>	- Prokaryotic Dynamic Programming Genefinding Algorithm
<b>PBVC</b>	- <i>Pyrodictum bahamense</i> var. <i>compressum</i>
<b>PEAR</b>	- Paired-end read merger
<b>pM</b>	- pico molar
<b>PO<sub>4</sub></b>	- Phosphate
<b>PSP</b>	- Paralytic Shellfish Poisoning
<b>PST</b>	- Paralytic Shellfish Toxin
<b>Rpm</b>	- revolution per minute
<b>rRNA</b>	- ribosomal ribonucleic acid
<b>rDNA</b>	- ribosomal deoxyribonucleic acid
<b>S</b>	- Second
<b>SRA</b>	- Sequence Read Archive
<b>SSU</b>	- Small subunit
<b>STX</b>	- Saxitoxin
<b>Sxt</b>	- Saxitoxin
<b>TD</b>	- Tagment DNA Buffer
<b>TBE</b>	- Tris Borate EDTA
<b>TE</b>	- Tris-HCl EDTA
<b>Tris-HCl</b>	- Tris (hydroxymethyl) aminomethane hydrochloride
<b>U</b>	- Unit
<b>UV</b>	- Ultra violet
<b>V</b>	- Variable region
<b>v/v</b>	- Volume over volume
<b>w/v</b>	- Weight over volume
<b>WGS</b>	- Whole genome sequencing

## LIST OF APPENDICES

	Pages
Appendix A: f/2 Medium Recipe	144
Appendix B: List of software used for the data analysis	146
Appendix C: Cell count of the triplicates cultures of PBVC from day 0 to 30, the cumulative average and the standard deviation	147
Appendix D: Gram staining result of 74 bacterial isolates viewed under 20 X magnification, 40 X magnification (* <sup>1</sup> ), and 100 X magnification (* <sup>2</sup> )	148
Appendix E: The DNA profiles (concentration and qualities of DNA) of 74 PBVC bacterial isolates measured using a nanoDrop 1000 UV-Vis spectrophotometer and their genomic DNA extraction gel photo viewed using 1% (w/v) agarose gel stained with ethidium bromine. Lane M: Lamda <i>HindIII</i> DNA ladder; Lanes with numbers: PBVC isolate codes.	158
Appendix F: PCR amplification of the 16S rRNA gene (V3-V4 region) for 74 bacterial isolates. Lane M: 100 bp DNA Ladder; Lanes with numbers: PBVC isolate codes	161
Appendix G: Strain ID, NCBI accession numbers, taxonomic name and sizes for the 74 isolates. The accession numbers obtained from deposited sequences in NCBI database	163
Appendix H: The frequency (in log scale) of the number of reads assigned at the Family level	165
Appendix I: Domain structure of the sxt-related proteins in PBVC088 compared to <i>Cylindrospermopsis raciborskii</i> T3 cyanobacterial sxt protein domain	166

# 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 Bacteroidetes (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.



UMS  
UNIVERSITI MALAYSIA SABAH

### 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.



UMS  
UNIVERSITI MALAYSIA SABAH