ABSTRACT

Soil bacterial diversity from three locations on maritime Antarctica was analyzed using 16S rRNA gene clone library construction and restriction fragment length polymorphism (RFLP) fingerprinting. Soil samples were from near Rothera Research Station in Rothera Point from Adelaide Island, from Viking Valley on north-eastern side of Mars Glacier from Alexander Island, and from Léonie Island on northern Marguerite Bay, Antarctic Peninsula. Five hundred and forty-eight clones were screened by RFLP and representatives of each phylotype were sequenced for identification. The phylotype sequences showed close relationship (i.e. ≥95% similarity) with bacterial divisions *Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Chloroflexi*, *Cyanobacteria*, *Firmicutes*, *Gemmatimonadetes*, *Proteobacteria*, *Verrucomicrobia* and *Unclassified Bacteria.* The least number of phylotypes were observed in Rothera Point soil sample (25) whereas the highest phylotype diversity belonged to Léonie Island (35). Certain phylotypes were exclusive to one site or two, whereas 33% of the phylotypes were shared by all clone libraries. Shannon diversity index (*H'*) revealed the highest bacterial diversity in Léonie Island (3.14) and lowest diversity in Rothera Point (2.93). The soil from Viking Valley showed high diversity (*H'*=3.09) comparable to that of the vegetated soil of Léonie Island despite the severity of its climate condition. There is an evident environmental influence on the pattern of biodiversity where the human-disturbed soil sample of Rothera Point revealed less bacterial diversity than the undisturbed soils of Léonie Island and Viking Valley.

ACKNOWLEDGEMENTS

I wish to begin with the name of God, who is the meaning of everything and without whom none of this would be. He is the one true light that I seek in darkness, and is the only one who keeps me sane throughout hard times. There is no possible way that I can ever show my gratitude and humbleness towards Him and for that I bow before Him alone and say: “al-Hamdulillah” (thanks to God).

In light of that, I would like to show my respect and appreciation to my supervisor, Prof. Dr. Irene Tan Kit Ping, for giving me the opportunity to work as one of the group members despite my lack of expertise in the field of laboratory work. This project has definitely launched the ship of my career into the ever-expanding ocean of science. I would like to give my sincere thanks for all the support and guidance she has granted me.

For a ship to get home safe, it requires a handful of staff and crew running about and keeping things together. Therefore, I’d like to thank the crew members who welcomed us aboard, our co-captain and consultant Mr. Chong Chun Wie for all his help during experimental work and data analysis, Ms. Goh Yuh Shan for constant advisory and a helping hand when supplies were running low, and all the other seniors who showed me around the place and helped me get started.

My regards also go to my loyal friends and lab-mates, Mrs. Maryam Karimian Oloumi and Ms. Elaheh Movahed for being there for me whenever I needed a little escape to paradise. Many thanks to both who are true friends and companions in good times and in bad, and for making this experience an enjoyable one indeed.

And last but certainly not least, to the two most important people in my life, I dedicate this to my parents for believing in me when I had no trust in myself and for pushing me beyond my boundaries. I can simply say thank you for all you have done and God bless you.

TABLE OF CONTENTS

CONTENT **PAGE**

ABSTRACT iii
ACKNOWLEDGMENTS iv
TABLE OF CONTENTS v
LIST OF FIGURES viii
LIST OF TABLES x
ABBREVIATIONS xi

CHAPTER 1. INTRODUCTION **1**

1.1 Objectives 3

CHAPTER 2. LITERATURE REVIEW **4**

2.1 Antarctica 4
2.1.1 History 4
2.1.2 Geography and climate conditions 5
2.1.3 Flora and fauna 7
2.1.4 Microbial diversity 8
2.2 Biodiversity; an overview 9
2.3 Significance of biodiversity 10
2.3.1 Story of evolution 10
2.3.2 Circle of life 11
2.3.3 Biotechnology perspective 13
2.4 Scale estimation 14
2.4.1 Culture-dependent techniques 15
2.4.2 Culture-independent techniques 16
2.4.2.1 DNA Extraction 16
2.4.2.2 Pursuing the gene 18
2.4.2.3 Primers and polymerase chain reaction (PCR) 19
2.4.2.4 Separation via cloning 21

2.4.2.5 Molecular fingerprinting 23
2.4.2.6 DNA sequencing 25
2.4.2.7 Identification 26
2.5 Biodiversity assessment 29
2.5.1 Diversity indices 29
2.5.2 Species richness estimation 30
2.6 Final remark 33

CHAPTER 3. METHODOLOGY 35

3.1 Study sites 35
3.2 Soil DNA extraction 37
3.2.1 DNA concentration 37
3.3 PCR amplification of 16S rRNA gene 37
3.4 Agarose gel electrophoresis 39
3.5 PCR product purification 40
3.6 16S rRNA gene cloning 41
3.6.1 Media preparation 42
3.6.1.1 Selective Luria Bertani (LB) agar 42
3.6.1.2 Super Optimal Broth with Catabolite Repression (SOC) medium 42
3.6.2 Ligation 43
3.6.2.1 Insert:vector ratio calculation 43
3.6.2.2 Cloning PCR product with pGEM®-T Easy Vector 43
3.6.3 Transformation 44
3.6.3.1 Cold $CaCl_2$ treatment 44
3.6.3.2 Heat shock 45
3.6.3.3 Pre-incubation 45
3.6.3.4 Plating 45
3.6.4 Screening 46
3.7 Detection of positive insert 46
3.7.1 SP6 and T7 promoter PCR amplification 46
3.7.2 Glycerol stock 47
3.8 Restriction fragment length polymorphism (RFLP) 47
3.8.1 Restriction enzyme digestion with *HaeIII* and *HhaI* 47
3.8.2 RFLP analysis 48

3.9 Sequencing 49
3.10 Data analysis 49
3.10.1 Phylogenetic analysis 49
3.10.2 Statistical analysis 51

CHAPTER 4. RESULTS 52

4.1 Genomic DNA concentration and purity 52
4.2 Primary PCR reaction 52
4.2.1 Template 53
4.3 PCR purification 54
4.4 Cloning 55
4.5 RFLP analysis 58
4.6 Sequencing 60
4.7 Data analysis 61
4.7.1 16S rRNA gene sequences 61
4.7.2 Phylogenetic analysis 61
4.7.3 Statistical analysis 64
4.7.3.1 Observed number of phylotypes 64
4.7.3.2 Biodiversity indices 64
4.7.3.3 Accumulation curves and richness estimators 65

CHAPTER 5. DISCUSSION 69

5.1 Conclusion 77

APPENDICES 78

Appendix A 78
Appendix B 79
Appendix C 81
Appendix D 82
Appendix E 84

REFERENCES 86

LIST OF FIGURES

FIGURES	PAGE
Figure 2.1 Map of Antarctica continent (retrieved from: http://www.map-of-antarctica.us/, on May 3, 2010)	5
Figure 3.1 a. Map of Antarctic Peninsula showing the two main Islands where study sites are located b. Map of Rothera Point on Adelaide Island and Léonie Island (obtained from Convey & Smith, 1997) c. Map of Viking Valley on Alexander Island (Convey & Smith, 1997)	36
Figure 3.2 5µl HyperLadder I/lane, 1% molecular biology grade agarose in 1× TAE stained with Ethidium Bromide (retrieved from: http://www.bioline.com/h_ladderguide.asp, on May 19, 2010)	39
Figure 3.3 pGEM®-T Easy Vector map and sequence reference points showing multiple cloning sequences (Technical Manual for pGEM®-T and pGEM®-T Easy Vector Systems, retrieved from: http://www.promega.com/tbs/tm042/tm042.pdf, on May 30, 2010)	41
Figure 4.1 Primary PCR amplification of RO site DNA sample a. Resulting bands after 50× (left) and 20× dilutions (right). The lanes are marked according to the volume of template used; no band is observed in negative control b. Purified DNA	54
Figure 4.2 Blue/white colony screening of plates of RO cloning session (from left to right, top row transformation and negative control plates, bottom row sample and positive control plates)	56
Figure 4.3 a. the promoter and MCS of pGEM®-T Easy Vector (Technical Manual for pGEM®-T and pGEM®-T Easy Vector Systems, retrieved from: http://www.promega.com/tbs/tm042/tm042.pdf, on May 30, 2010)	57
Figure 4.3 b. Colony PCR of cloned cells from RO cloning session; all bands showing the correct size except one (pointed out with arrow)	58
Figure 4.4 An overview of RFLP analysis corresponding to ROI clone library with examples of some of the patterns. Lambda DNA (image to the right) was used as a control for monitoring enzyme digestion.	59
Figure 4.5 PCR amplification of 16S rRNA gene; the sample showing a second band at ~300bp (pointed out by arrow) was discarded.	60
Figure 4.6 Bar charts of clone library composition at phylum level for Rothera (RO), Viking Valley (VV), and Léonie Island (LE) samples	62
Figure 4.7 Log abundance plots of observed phylotypes considering their relative abundance in each clone library	66

Figure 4.8 a. Phylotype accumulation curves (collector's curves) for different clone libraries from study sites b. phylotype richness corresponding to each clone library estimated by *Chao1* unbiased formula 67

Figure 4.9 Phylotype level non-metric multidimensional scaling (nMDS) of clone library data from three different sites in Antarctica 68

Figure 1S Example of sequence chromatogram obtained from clone ROI-2 81

Figure 2S Unrooted Neighbor-Joining phylogenetic tree of 16S rRNA gene sequences using Maximum Composite Likelihood method. Representatives of phylotypes along with the occurrence in respective sites are shown. GenBank accession numbers of supplementary sequences are presented in square brackets. Bootstrap analysis with 500 repetitions was calculated and values above 50% are shown. 84

LIST OF TABLES

TABLES **PAGE**

Table 3.1 Study sites specifications and soil characteristic 35

Table 3.2 Master mix composition for different number of PCR reactions 38

Table 3.3 SOC media chemical components in 1000ml LB broth 43

Table 3.4 Ligation reaction set up for cloning PCR product with pGEM$^{®}$-T Easy Vector System 44

Table 3.5 Restriction enzyme reaction preparation 47

Table 4.1 Genomic DNA concentration of soils collected from Rothera (RO), Viking Valley (VV) and Léonie Island (LE) 52

Table 4.2 Purified DNA concentration post PCR amplification and calculated amount of DNA required for ligation reaction for each cloning session (duplicates carried out for each site indicated by Roman numerals) 55

Table 4.3 Colony count of duplicate cloning sessions, number of white colonies and blue colonies for both sample and positive control plates 57

Table 4.4 Number of clones from soil clone libraries of three different sites in Antarctica, number of unique and community-dominant (CD) phylotypes, percentage of clones in CD phylotypes and sampling coverage for each clone library 60

Table 4.5 Diversity indices calculated for soil clone libraries 65

Table 1S Advantages and disadvantages of some molecular-based methods to study soil microbial diversity (Kirk *et al.*, 2004) 78

Table 2S Phylotypes and their closest sequence match from GeneBank (retrieved from: http://www.ncbi.nlm.nih.gov/BLAST, on June 15, 2010) 79

Table 3S Phylotype classifications (retrieved from: http://rdp.cme.msu.edu/index.jsp, on June 22, 2010) 82

ABBREVIATIONS

%	: percent
=	: equal to
≥	: greater than or equal to
≤	: smaller than or equal to
℃	: degree Celsius
µg	: microgram
µl	: microlitre
µM	: micromolar
A_{260}/A_{230}	: ratio of UV absorbance at 260nm and 230nm
A_{260}/A_{280}	: ratio of UV absorbance at 260nm and 280nm
ARDRA	: amplified ribosomal DNA restriction analysis
ATS	: Antarctic Treaty System
BAS	: British Antarctic Survey
BLAST	: basic local alignment search tool
bp	: base pairs
$CaCl_2$	: calcium chloride
cAMP	: cyclic adenosine monophosphate
CAP	: catabolite activator protein
CD	: community-dominant
cm	: centimetre
D	: Simpson’s diversity index
DGGE	: denaturing gradient gel electrophoresis
dH_2O	: distilled water
DNA	: deoxyribonucleic acid
dNTP	: deoxyribonucleoside triphosphate
dsDNA	: double-stranded deoxyribonucleic acid
EDTA	: ethylenediaminetetraacetate acid
E. coli	: *Escherichia coli*
ET	: extra-terrestrial
EtBr	: ethidium bromide
F	: forward
g	: gram

GPS : Global Positioning System

H' : Shannon's diversity index

*Hae*III : *Haemophilus aegypticus*III

HCl : hydrochloric acid

*Hha*I : *Haemophilus haemolyticus*I

IPTG : isopropylthiogalactose

J' : Pielou's evenness

KCl : potassium chloride

L : litre

LB : Luria Bertani

LE : Léonie Island

M : molar

MCS : multiple cloning site

mg : milligram

$MgCl_2$: magnesium chloride

min : minute

ml : millilitre

mM : millimolar

NaCl : sodium chloride

NJ : neighbour-joining

nm : nanometre

nMDS : non-metric multidimensional scaling

OD : optical density

OTU : operational taxonomic unit

PCR : polymerase chain reaction

PSI : percentage sequence identity

R : reverse

rDNA : ribosomal deoxyribonucleic acid

RDP : Ribosomal Database Project

RE : restriction enzyme

RFLP : restriction fragment length polymorphism

RNA : ribonucleic acid

RO : Rothera Point

rpm : revolutions per minute

rRNA : ribosomal ribonucleic acid

sec	: second
SOC	: super optimal broth with catabolite repression
SSCP	: single strand conformation polymorphism
ssDNA	: single-stranded deoxyribonucleic acid
ssu	: small subunit
TAE	: tris acetate ethylenediaminetetraacetate acid
Taq	: *Thermus aquaticus*
TGGE	: temperature gradient gel electrophoresis
UV	: ultraviolet
V	: volt
VV	: Viking Valley
w/v	: weight per volume
X-Gal	: 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside