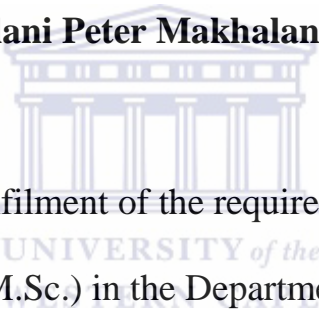


The diversity of key anabolic genes in Antarctic hypolithons

By

Thulani Peter Makhalanyane

The logo of the University of the Western Cape, featuring a classical building with columns and a pediment, with the text 'UNIVERSITY of the WESTERN CAPE' below it.

Submitted in partial fulfilment of the requirements for the degree of
Magister Scientiae (M.Sc.) in the Department of Biotechnology,
University of the Western Cape

Supervisor: Prof. D.A. Cowan

Co-Supervisor: Dr. IM Tuffin

November 2009

Declaration

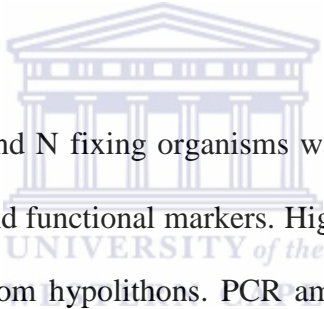
I declare that ‘The Diversity of key anabolic genes in Antarctic hypolithons’ is my own work, that it has not been submitted for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged by complete references.

Thulani Makhwanyane



Abstract

Antarctica is known for its pristine environments. A variety of unsuitable environmental conditions were once thought to render the continent unsuitable for sustaining life. However, metagenomic data have revealed a wealth of species diversity in a range of biotopes. Hypolithons, photosynthetic communities which live under translucent rocks in climatically extreme environments, are an important input source for both carbon (C) and nitrogen (N) in this hyperarid desert environment. Microbial contribution to biogeochemical cycling resulting in fixation of both C and N remains poorly understood. Moreover, there is a reported close interplay between both cycles, with nitrogen being reported to be a limiting factor in carbon assimilation.



In this study the diversity of C and N fixing organisms was investigated by using the *cbbL* and *nifH* genes as phylogenetic and functional markers. High Molecular weight metagenomic DNA and RNA was extracted from hypolithons. PCR amplification was carried out using *cbbL* (800 bp for red-like, 1,100 bp for green-like) and *nifH* (360 bp) gene specific primers. Resultant PCR products were used to construct libraries which were screened for correct sized inserts. Restriction Fragment Length Polymorphism (RFLP) was used to de-replicate clones prior to sequencing. Phylogenetic positions from both clone libraries were established by aligning nucleotide sequences and constructing similarity trees using NJ clustering methods.

BLASTn results indicated the presence of previously uncultured organisms which contain *cbbL* and *nifH* genes. BLASTn results were characterized by low percentages of maximum identity (typically <95%), a potential indicator of novel taxa. Sequences from respective libraries clustered with cyanobacteria such as *Nostoc*, *Scytonema*, and *Tolypothrix* and α -, β -,

and γ -Proteobacteria such as *Azotobacter*, *Agrobacterium* and *Mesorhizobium*. Generally sequence results indicate a largely homogenous, being dominated by specific taxa. Each group may contain potential keystone species, essential for both biogeochemical cycling in oligotrophic environment.



Publications originating from this Thesis

Poster/Presentations

1. **Makhalanyane, T. P., Mutondo, M., Tuffin, I. M. & Cowan, D. A. (2009).** Diversity of “Red-like” *cbbL* genes in Antarctic hypolithons. *SANAP Symposium 2009*. Cape Town: South Africa. **Poster**
2. **Mutondo, M., Makhalanyane, T. P., Tuffin, I. M. & Cowan, D. A. (2009).** Metatranscriptomic analysis of Antarctic hypolithic communities. *SASM 2009*. Durban: South Africa. **Poster**
3. **Makhalanyane, T. P., Mutondo, M., Tuffin, I. M. & Cowan, D. A. (2009).** Diversity and expression of key anabolic genes in Antarctic hypolithons. *SASM 2009*. Durban: South Africa. **Talk**
4. **Makhalanyane, T. P., Mutondo, M., Tuffin, I. M. & Cowan, D. A. (2009).** Diversity and expression of key anabolic genes in Antarctic hypolithons. *UWC Open Day 2009*. Cape Town: South Africa. **Talk**

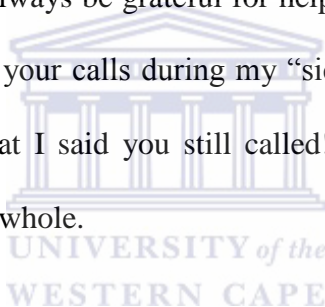
Publications

5. **Cowan, D. A., Sohm, J. A., Makhalanyane, T., Capone, D. G., Green, T. G. A. & Cary, S. C.** Hypolithic communities: a hidden nitrogen source in Antarctic desert soils. *Science*, **In Preparation**.

Acknowledgements

I would like to thank the **NRF** and **SANAP** for funding which enabled me to pursue this project.

I would like to thank my supervisor, **Professor Don Cowan** for giving me an opportunity to work on an inspired, novel research project for my MSc. I am grateful for the opportunity to work in your laboratory, and your guidance throughout this project. A word of thanks to **Dr. Marla Tuffin** for all the belief you had in my abilities as a student and for your unending support. To **Dr. Heide Goodman** - I am grateful for the caring role that you have played over the course of this project. I will always be grateful for helping to set me up in Cape Town. I will especially always remember your calls during my “sick leave” (even though you could probably not hear a word of what I said you still called!). I would also like to thank the department of biotechnology as a whole.



I am deeply indebted to the following people who read drafts of this thesis; Dr Mark Paul Taylor (Ch 1), Dr Walter Sanyika (Ch 1 and 3) Dr Cal Heath (Ch 1), Dr Bronwyn Kirby (Ch 3 and 4), Dr Moola Mutondo (Ch 2), Dr Francesca Stomeo (Ch 1), Dr Inonge Mulako (Ch 1), Prof Adolf Zeze (Ch 1 and 2), Mr. Victor Muleya (Ch 1), Mr William Mavengere (Ch 1, 2 and 3) and Mr Lonnie van Zyl (Ch 1, 3). Your contribution towards helping me put my thoughts down on paper, in the most non-journalistic, scientific manner is highly appreciated.

I am also appreciative of everyone who has helped teach me some of the techniques that I have learned over the course of my MSc; Prof Kathy Coyne (University of Delaware, RNA extractions), Dr Nuraan Khan (DGGE), Dr Walter Sanyika and Mr J Lako (Too many to mention!).

I am grateful to a number of people at IMBM who have been kind to me over the course of this project. **Ncebakazi, Hanna**, Clive, Farzaana, Dean, Natasha, William, Andrew, Claude, Babele, Layla-Lusinda, Oratilwe, Victor, Tshifhiwa, Billy, Khomotso, thank you for being such pleasant people to be around (most of the time ☺). To everyone else who has contributed to the day to day running of IMBM, namely, Nazneen, Auntie Ana and Fungai – your role is very much appreciated!

To **Andiswa** thanks for keeping me sane. The past year would not have been the same without you.

Thank God for everyone else who has contributed to the successful completion of this project. It takes a village to produce a thesis!



Dedication

To papa and mama, your belief in me has allowed me to reach this point. Your continuous encouragement and sustenance has enabled me to fulfil my dream.

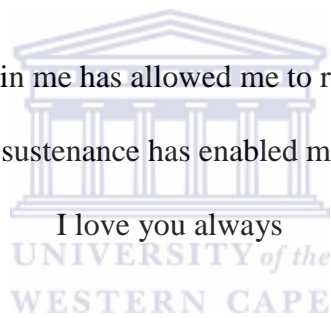
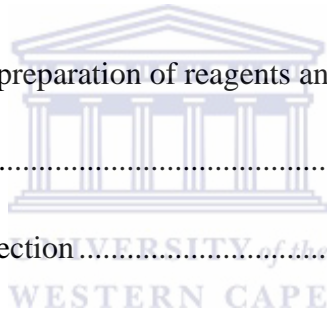


Table of Contents

Declaration.....	ii
Abstract.....	iii
Publications originating from this Thesis	v
Acknowledgements.....	vi
Dedication.....	viii
Table of Contents.....	ix
List of Abbreviations	xix
Chapter 1: Literature Review.....	1
1.1 Antarctica.....	1
1.1.1 Antarctic Dry Valleys	2
1.1.2 Climate.....	4
1.1.3 Biotopes	5
1.2 Hypolithons.....	7
1.3 Global and Regional Carbon Ecology	9
1.4 Biological Carbon Cycle.....	10
1.4.1 The Calvin-Benson-Bassham (CBB) Cycle.....	12
1.4.2 Ribulose 1, 5-bisphosphate carboxylase/oxygenase (RubisCO)	13
1.4.3 Diversity of potential CO ₂ fixing bacteria in soil	16
1.5 Nitrogen Fixation.....	16



1.5.1 Diazotrophs	18
1.5.2 Nitrogenase	19
1.5.2.1 Dinitrogenase and Dinitrogenase Reductase	20
1.5.3 Major <i>nifH</i> gene clusters	20
1.5.4 Diversity of potential nitrogen fixing microorganisms.....	21
1.6 Aims and Objectives	22
1.5.1 Aims.....	23
1.5.2 Objectives	23
Chapter 2: Materials and Methods.....	24
2.1 General procedures for the preparation of reagents and glassware for RNA work.....	24
2.2 Media	24
2.3 Environmental Sample collection.....	27
2.4 DNA/RNA Extractions	27
2.4.1 Modified Miller Protocol	27
2.4.2 DNA extraction from bacterial cultures.....	28
2.4.3 RNA extraction using MoBio Soil RNA kit.....	28
2.4.4 Modified DNA/RNA protocol using CTAB lysis buffer.....	29
2.4.4 Purification using Polyvinylpyrrolidone (PVPP) minicolumns	30
2.5 Bacterial Strains	30
2.6 Polymerase Chain Reaction	31
2.7 Analytical procedures	33



2.7.1 Agarose gel electrophoresis	33
2.7.2 Fluorometry (Qubit™).....	33
2.7.3 Spectrophotometry	34
2.8 Gel Extraction and Purification	35
2.9 Denaturing Gradient Gel Electrophoresis (DGGE)	35
2.10 Construction of Clone libraries	36
2.10.1 Preparation of <i>Escherichia coli</i> (<i>E. coli</i>) competent cells.....	36
2.10.2 Ligation of PCR products	37
2.10.3 Transformation of competent <i>E. coli</i> Gene Hogg strain	37
2.11 Screening of clone libraries	38
2.11.1 Colony PCR	38
2.11.2 Plasmid DNA extraction	39
2.11.3 Restriction Fragment Length Polymorphism (RFLP).....	39
2.11.3.1 Analysis of RFLP patterns	39
2.12 Sequencing and Sequence analysis	40
2.12.1 Sequencing reactions	40
2.12.2 Sequence Analysis and Construction of phylogenetic trees	40
Chapter 3: Diazotrophic diversity and expression in an Antarctic hypolithon.....	41
3.1 Introduction.....	41
3.2 DNA and RNA Extraction	43
3.2.1 Simultaneous Extraction using CTAB lysis buffer.....	43

3.2.3 MoBio Soil RNA kit	45
3.3 Expression of <i>nifH</i> genes	46
3.3.1 DNase treatment of simultaneous extractions.....	46
3.4 Diversity of <i>nifH</i> genes	48
3.4.1 RNase treatment of simultaneous extractions.....	48
3.4.2 PCR amplification of the 16S rRNA gene with DGGE primers	49
3.4.3 DGGE	50
3.4.4 PCR with <i>nifH</i> primers.....	55
3.4.5 <i>nifH</i> Clone library	56
3.4.6 RFLP analysis of clones.....	57
3.4.6.1 Richness Estimation of RFLP groups	58
3.4.7 Phylogenetic Analysis.....	61
3.5 Discussion.....	65
Chapter 4: Phylogenetic diversity of autotrophic carbon sequestering genes in Antarctic hypolith communities.....	67
4.1 Introduction.....	67
4.2 Results and Discussion	69
4.2.1 DNA extraction from environmental samples	69
4.2.2 Amplification of <i>cbbL</i> genes	69
4.2.3 <i>cbbL</i> Clone library	70
4.2.3.1 <i>cbbL</i> Red-like.....	70
4.2.3.2 <i>cbbL</i> Green-like	71



4.2.4	Diversity indices of clone libraries	72
4.2.4.1	<i>cbbL</i> Red-like.....	72
4.3	Phylogenetic analysis based on <i>cbbL</i> sequence data	75
4.3.1	<i>cbbL</i> red-like	75
4.3.2	<i>cbbL</i> green-like	78
4.4	Discussion	80
Chapter 5: Conclusions and Future Prospects		82
Appendices.....		87
References.....		110



List of Tables

Table 1: Summary of the properties of different forms of RubisCO (Tabita <i>et al.</i> , 2008).....	13
Table 2: Composition of major nitrogenase gene clusters (Zehr <i>et al.</i> , 2003).....	20
Table 3: Primer combinations, cycling conditions and specificity of primers used in this study.....	32
Table 4: Preparation of assay tubes was carried out as follows.....	34
Table 5: Richness Estimation Prediction model for <i>nifH</i> Clone library	59
Table 6: Richness Estimation Prediction model for <i>cbbL</i> red-like clone library.	73



List of Figures

Figure 1: Map of continental Antarctica indicating key research sites (Cowan & Ah Tow, 2004).	2
Figure 2: Potentially significant drivers for microbial diversity in Antarctic Dry Valley biotopes.	3
Figure 3: Aerial photograph of the Wright Valley, Antarctica. A range of biotopes can be observed from the photograph (Photo courtesy of Don Cowan).	6
Figure 4: Translucent quartz rock with classic underside colonization by hypolithic communities (Photo courtesy of Don Cowan).	8
Figure 5: The <i>cbbL</i> gene as a marker for CO ₂ fixation is involved in the first rate limiting step in the CBB cycle.	11
Figure 6: The Calvin-Benson-Bassham cycle (Atomi, 2002).	12
Figure 7: Cartoon representative structures of different Forms of RubisCO (Tabita, 1999)...	14
Figure 8: Nitrogen cycling and genes involved in biogeochemical N cycling. The <i>nifH</i> gene represents the rate limiting step in the nitrogen cycle.	17
Figure 9: Feedback pathways Interaction between C and N cycle. There is a close relationship between mineral nitrogen and atmospheric CO ₂ fixation through the action of plants. In Antarctic environments microorganisms are expected to fulfill the role normally taken by plants (Thornton <i>et al.</i> , 2009).	23
Figure 10: Simultaneous extractions of DNA and RNA from hypolith samples, denoted Miers Valley Hypolith (MVH)). Lane 1: RNA Marker (Fermentas) Lane 2: MVH 3A; Lane 3: MVH4A. 23S and 16S rRNA bands are only faintly visible. (Electrophoresis on 1% agarose gel); Lane 4: Blank.	44

Figure 11: Simultaneous DNA/RNA extraction samples. Lane 1: Marker (IMBM marker), Lane 3: MVH 3A, Lane 5: MVH 4A, Lane 7: MVH 11A, Lane 9: MVH 12A (1% agarose gel electrophoresis). Lanes 2, 4, 6, and 8: Blank.44

Figure 12: Electrophoresis of RNA extracted using the MoBio kit.....45

Figure 13: 4 DNase I treatment of Simultaneous extracts (DNA and RNA).46

Figure 14: 16S rRNA PCR amplicons on DNase I treated simultaneous extracts indicating amplification at various concentrations. Lane 1: Marker (IMBM Marker), Lane 2 to Lane 6: different concentrations of template RNA (ranging from 10 ng/ul to 50 ng/ul). Amplification was most efficient with 30 ng/ul of template. Lane 7: Negative control (No template) (1% agarose gel electrophoresis).47

Figure 15: RNase treatment of DNA/RNA extracts to remove RNA.....49

Figure 16: PCR amplification of bacterial 16S rRNA using primers 341 F-GC and 534R on 1% agarose gel. Lane 1: Biolabs 1kb Marker, Lane 2: Blank; Lane 3: *E. coli* Positive Control, Lane 4 -6: Various hypolith soil samples. Lane 7: Negative control.....50

Figure 17: DGGE Analysis of Antarctic Dry Valley hypolith samples, open soil and non-translucent rock soil. Duplicate extractions were carried out for each sample before separation on DGGE. Lane 1 DGGE Marker, Lane 2 and 3: MVH 3A, Lane 4 and 5: MVH 4A, Lane 6 and 7: MVH 5A, Lane 8 and 9: MVH 9A, Lane 10 and 11: MVH 11A, Lane 12 and 13: MVH 12A, Lane 14 and 15, MVH 3C (Open soil), Lane 16 and 17: MVH 3B (non-translucent rock), Lane 18 and 19: *Nostoc sp.* DNA.51

Figure 18: Gelcompar Screen-shot of (A) Similarity dendrogram generated from DGGE profile. (B) DGGE profile with electrophoresis peaks from bands. (C) Sample sorted according to hierarchical clustering from bands.53

Figure 19: MDS ordinations showing relationship between terrestrial habitats based on the presence of bacterial genospecies. Hypolithic communities are closely related to each other, as indicated by observed clusters.....54

Figure 20: Amplification of the *nifH* gene fragment from hypolithic samples. Lane 1: *Azotobacter vinelandii*, Lane 2: MVH 11A Lane 3: 100bp Marker (Fermentas) (1% Agarose gel of *nifH* amplified fragment)55

Figure 21: Colony PCR indicating the presence of *nifH* fragments in over 80% of screened clones56

Figure 22: Agarose gel electrophoresis of *HaeIII* RFLP patterns for clones obtained from M13 PCR products.....57

Figure 23: Library richness estimation. In a well sampled library there is concurrence in the shape of pane A and C with B and D.....60

Figure 24: The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 9.11 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base.....62

Figure 25: Agarose gel (1%) electrophoresis of DNA extractions. Lane 1: λ - Pst Marker, Lane 2 and Lane 3: MVH 3A, Lane 4 and 5: Control soil (UWC).....69

Figure 26: Agarose gel (1%) electrophoresis of amplified *cbbL* amplicons. (A) Lane 1: Marker (IMBM marker), Lane 2: Negative control, Lane 3: *Bacillus licheniformis* (Positive control), (B) Lane 1: λ Pst Marker:, Lane 2: Positive Control Lane 2: MVH 3A sample Lane 3: Negative control..... 70

Figure 27: 2.5% Agarose gel of *cbbL* red-like RFLP digests with HaeIII. Lane 1: 100bp DNA ladder (Fermentas), Lane 2-25: Selected M13 amplicons. 71

Figure 28: 2.5% Agarose gel of random amplicons digested with Alu I. Lane 1: IMBM lab marker; Lanes 2-25: Selected restriction digests products. 72

Figure 29: Library richness estimation for *cbbL* red-like libraries. In a well sampled library there is concurrence in the shape of panel A and C with B and D (Kemp & Aller, 2004).... 74

Figure 30: The evolutionary history was inferred using the Neighbor-Joining method. evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total of 371 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4. 77

Figure 31: The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 2.78597539 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons (Pairwise deletion option). There were a total of 784 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4. 78

List of Abbreviations

$(\text{NH}_4)_2\text{SO}_4$	ammonium sulphate
$^{\circ}\text{C}$	degrees Celsius
μg	microgram
μl	microlitre
μM	micromolar
ATP	adenosine triphosphate
bp	base pairs
BSA	bovine serum albumin
CaCl_2	calcium chloride
CBB	Calvin-Benson-Bassham
cfu	colony forming units
cm	centimetre
CTAB	cetyl trimethyl ammonium bromide
DGGE	denaturing gradient gel electrophoresis
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
EDTA	ethylene diamine tetraacetic acid
H'	Shannon diversity index
HCl	hydrochloric acid
hrs	hours
IPTG	isopropyl β -D-thiogalactosidase
KCl	potassium chloride
km	kilometre

l	litre
LB	Luria Bertoni medium
LB-amp	Luria Bertoni medium containing ampicillin
m	metre
M	molar
MDS	multi-dimensional scaling
mg	milligram
MgCl ₂	magnesium chloride
MgSO ₄	magnesium sulphate
min	minutes
ml	millilitre
mm	millimetre
mM	millimolar
MnCl ₂	manganese chloride
NaCl	sodium chloride
NaH ₂ PO ₄	sodium dihydrogen orthophosphate
NaOH	sodium hydroxide
<i>nifH</i>	nitrogenase reductase
ng	nanogram
nm	nanometre
OD ₆₀₀	optical density at 600nm
PCR	polymerase chain reaction
rpm	revolutions per minute
RFLP	restriction fragment length polymorphism
rRNA	ribosomal oxyribonucleic acid



s	seconds
SDS	sodium dodecyl sulphate
SSU	small subunit
TAE	tris acetic acid EDTA
UV	ultraviolet
v/v	volume per volume
w/v	weight per volume
X-gal	5-bromo-4-chloro-3-indolyl- β -D-galactosidase



Chapter 1: Literature Review

1.1 Antarctica

When men first landed on Antarctica in the beginning of the 19th century, the overwhelming motivating factor for human presence on the continent and in the surrounding seas could be summarized primarily as a search for knowledge and a quest for economic gain (Figure 1) (Committee on Antarctic Policy and Science *et al.*, 1993). Antarctica as a place has at all times been full of surprises. In the eighteenth century, Captain James Cook anticipated the discovery of a land of forests and pastures ready for migration; instead he discovered an enormous ice-covered and apparently useless continent (Committee on Principles of Environmental Stewardship for the Exploration and Study of Subglacial Environments & National Research Council, 2007). Yet, in spite of this, each year an increasing number of explorers and researchers flock to the continent enchanted by its seamless beauty and mystery.

Antarctica as a continent is very isolated, with the closest landmass being more than a thousand kilometres away (Figure 1). By human standards, Antarctica is the most inhospitable of the continents (Onofri *et al.*, 2004). As the southernmost continent, Antarctica offers some of the most extreme recorded environmental conditions on Earth (Figure 2) (Fell *et al.*, 2006).

The continent has been intensely studied. There are a range of reasons why researchers from various disciplines have sought to use the continent for their studies. Antarctica has the cleanest air on earth, allowing for air quality monitoring with a reliable baseline. Antarctica is the darkest place on Earth, which makes it ideal for astronomical research. The relatively

simple environment has been used as an analog for space environments and planetary extremes (Andersen *et al.*, 1990). For microbial ecologist Antarctica offers the opportunity to probe a range of questions, owing in part, to the relative simplicity of its various environments.

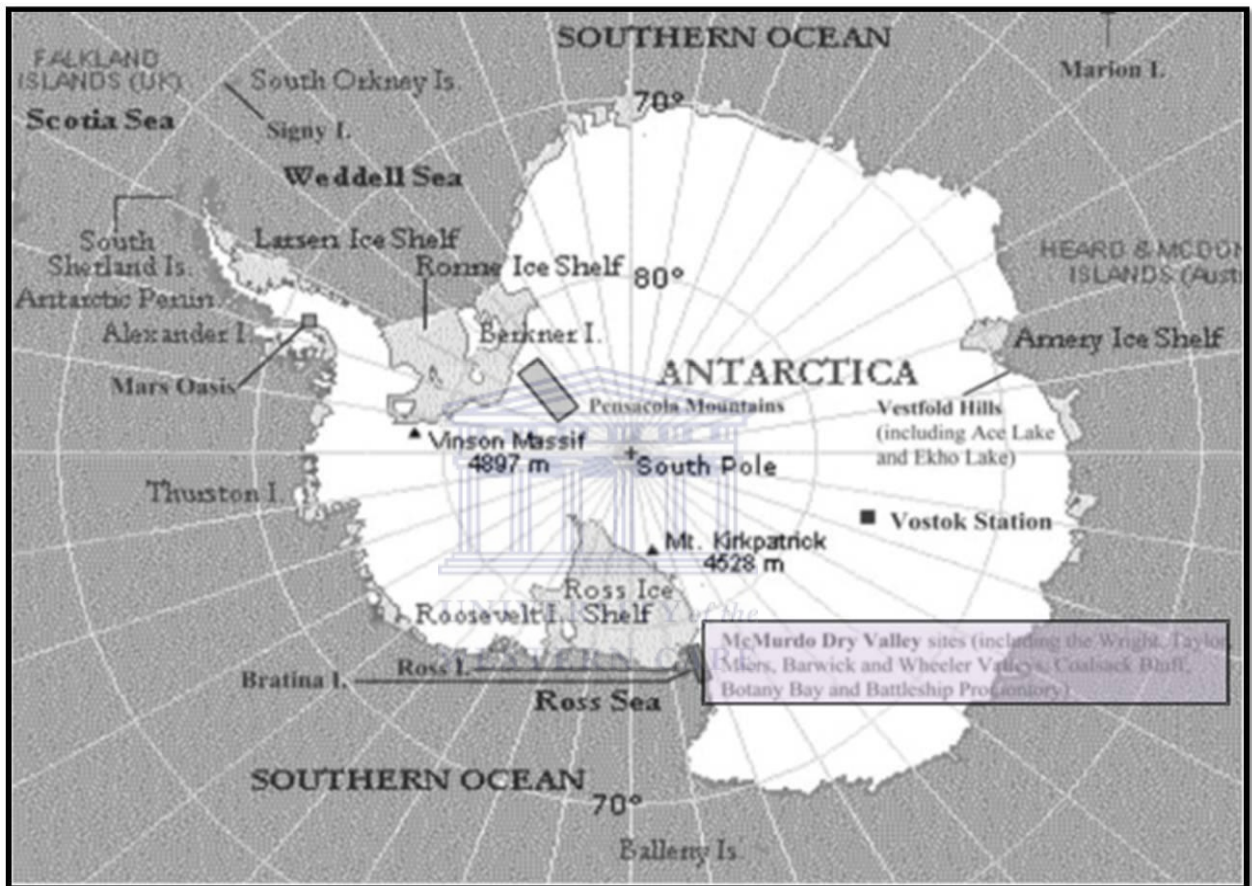


Figure 1: Map of continental Antarctica indicating key research sites (Cowan & Ah Tow, 2004).

1.1.1 Antarctic Dry Valleys

An estimated 98% of the 14.4 million kilometres square (km²) of the continent is ice covered. This leaves a mere 2% of ice-free regions. The Dry Valleys of Eastern Antarctica, which make up part of the ice-free regions, have some of the most extreme environmental conditions on earth (Figure 2) (Convey, 1997; Cowan & Ah Tow, 2004). The cumulative

effect of these climatic factors has resulted in an environment which is unfavourable for supporting life (Boyd & Boyd, 1963). Antarctica has, for a long time been considered to be a life-less continent of rock and ice (Weisburd, 1986). However suitable niches have been found and life does exist.

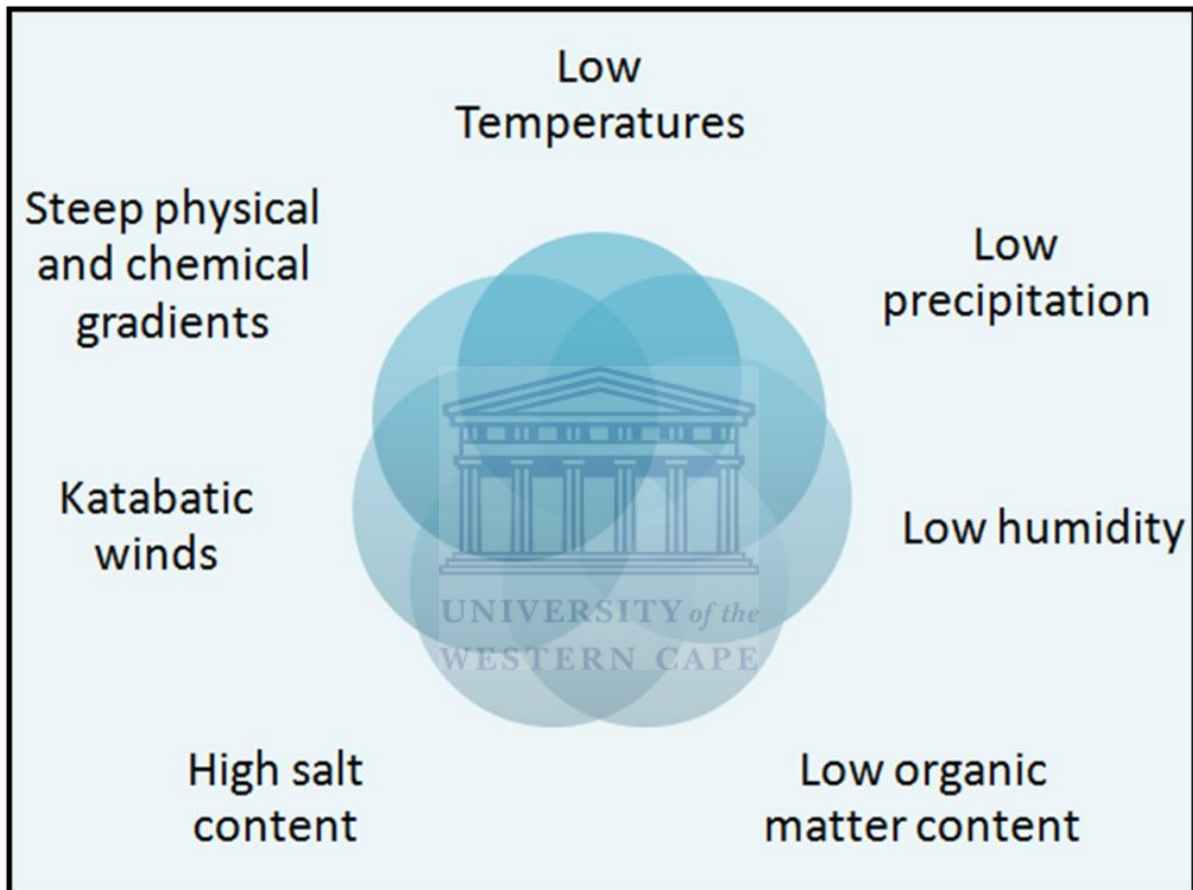
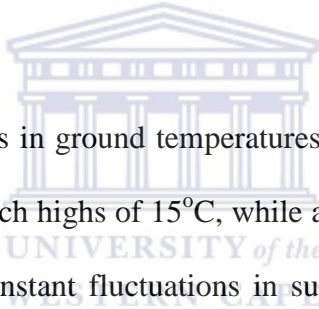


Figure 2: Potentially significant drivers for microbial diversity in Antarctic Dry Valley biotopes.

Several studies have characterized a group of invertebrates inhabiting the harsh Dry Valleys (Doran *et al.*, 2002; Stevens & Hogg, 2002; Treonis *et al.*, 1999). Conversely, other studies have reported an abundance of prokaryotes in these biotopes (Cowan *et al.*, 2002) much higher than previously reported.

1.1.2 Climate

The average annual temperatures in the Dry Valleys range from -25°C to -20°C (Ascaso & Wierzchos, 2002; Wynn-Williams, 1988), with wide fluctuations in summer temperatures. Summer air temperatures can range from -35°C to 3°C or higher (de la Torre *et al.*, 2003). The severity of low temperatures is varied depending on specific locations on the continent. This is due, in part, to approximate distances from the ocean, and moving from one latitudinal position to the next. Consequently, East Antarctica is colder than West Antarctica, due to its higher latitudinal elevation and proximity to the ocean. The Antarctic Peninsula has the most moderate climate; higher temperatures occur in January along the coast and average slightly below freezing.



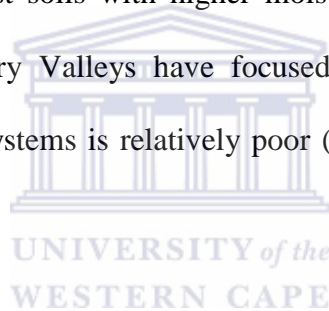
There are considerable differences in ground temperatures and air temperatures. In summer, the mean ground temperatures reach highs of 15°C , while air temperatures average 0°C (de la Torre *et al.*, 2003). There are constant fluctuations in surface ground temperatures, which arise as a consequence of prolonged direct sunlight or cloud cover. In summer, temperatures may change by as much as 15°C in minutes due to cloud cover (de la Torre *et al.*, 2003).

Climate studies carried out on Antarctica have reported varying continental trends. Doran and colleagues reported continental cooling through a 14 year Antarctic Dry Valley meteorological record over a 35 year period temperature compilation (Doran *et al.*, 2002). An opposing view by Jacka and colleagues reported a net continental warming (Jacka & Budd, 1998). A more recent study, which looked at meteorological trends over a 50 year period, found that temperatures vary throughout the continent (Turner *et al.*, 2005). A “rapid warming” is reported over the Antarctic Peninsula, with other regions reported to be more inconsistent (Turner *et al.*, 2005). This finding suggests that temperature increases may be a

threat in Antarctic environments. This would be consistent with reports that The Dry Valleys are among the most threatened environments on earth due to climate changes (Cowan & Ah Tow, 2004; Doran *et al.*, 2002).

1.1.3 Biotopes

Biotopes are areas of uniformity in environmental conditions and distribution of animal and plant life. The Dry Valley biotopes include both aquatic and terrestrial ecosystems (Figure 3). The aquatic ecosystem consists of glaciers, streams and lakes (both freshwater and saline). The terrestrial ecosystem consists of exposed soils, which range from barren mineral soils with little water content to moist soils with higher moisture content (Cowan & Ah Tow, 2004). Initial studies on the Dry Valleys have focused on aquatic ecosystems and the understanding of terrestrial ecosystems is relatively poor (Convey & Stevens, 2007; Cowan & Ah Tow, 2004).



Dry Valley soils have been described as the most barren soils found in the ice-free regions. These soils are highly concentrated in salts such as sodium, calcium, magnesium, chloride, sulphate and nitrate (Horowitz & Hubbard, 1972). This creates a steep concentration gradient which may create an impediment for microbial growth (Cowan & Ah Tow, 2004; Moorhead *et al.*, 1999). The arid soils also contain low concentrations of organic material (<0.1% of dry soil weight) which may further impede the growth of microbial communities (Cameron, 1971). This is in contrast to rock associated communities such as those inhabited by cosmoliths which often support microbial organic material visible to the naked eye (Melody *et al.*, 2001).

The warmer Antarctic peninsula terrestrial biotopes are populated by lower plants, small invertebrates and microorganisms (Convey & Stevens, 2007). Microorganisms belonging to various genera such as *Achromobacter*, *Arthrobacter*, *Bacillus*, *Corynebacterium*, *Flavobacterium*, *Micrococcus*, *Planococcus*, *Pseudomonas*, *Streptomyces*, and *Nocardia* have previously been isolated from the more extreme Dry Valley soils (Aislabie *et al.*, 2006; Babalola *et al.*, 2009; Pearce *et al.*, 2009; Smith *et al.*, 2006).

The ground surface in the Dry Valleys is comprised of dry permafrost that supports a surface terrain characterized by sandstone, quartz and granitic rocks (Pointing *et al.*, 2009). Rock (lithic) associated communities are important in terrestrial ground surface ecosystems. Studies have shown that lithic communities contribute to the existing organic pools through primary production (Moorhead *et al.*, 1999).

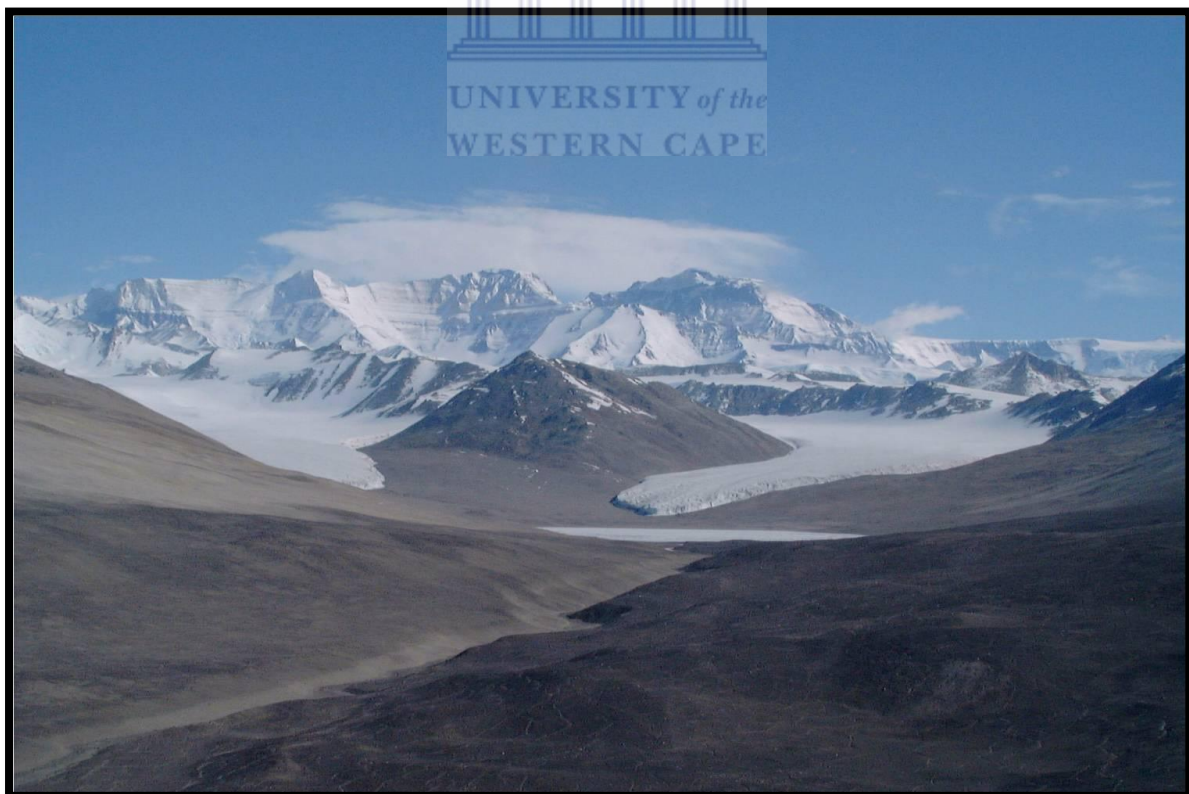
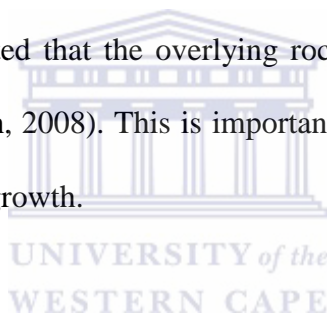


Figure 3: Aerial photograph of the Wright Valley, Antarctica. A range of biotopes can be observed from the photograph (Photo courtesy of Don Cowan).

1.2 Hypolithons

Microorganisms are able to adapt when faced with extreme environmental conditions. Hyperarid deserts, both hot and cold, offer examples of adaptation in microbial life (Pointing *et al.*, 2009). In these environments, microorganisms (called hypoliths) are forced to grow at the bottom of translucent rocks as a way of avoiding environmental extremes (Figure 4) (Cowan & Ah Tow, 2004; Warren-Rhodes *et al.*, 2006). Hypolithons represent unique “micro-biotope” where life thrives despite of environmental extremes.

In the Antarctic, the ability of hypolithons to tolerate extreme aridness may be an important trait that has allowed them to survive in conditions with low moisture (Warren-Rhodes *et al.*, 2006). A recent study has reported that the overlying rock of a hypolithon may provide a thermal buffer to hypoliths (Khan, 2008). This is important in hyperarid deserts where water is a limiting factor for microbial growth.



The quartz rock enables light to penetrate to the underside. The degree of light penetration has been shown to be sufficient to allow photosynthesis (Cockell & Stokes, 2004). Additionally, the translucent rock surface of the hypolithons provides a suitable habitat by trapping moisture essential for photosynthesis (Cowan & Ah Tow, 2004)



Figure 4: Translucent quartz rock with classic underside colonization by hypolithic communities (Photo courtesy of Don Cowan).

In oligotrophic deserts such as the Dry Valleys, hypolithons are potentially a significant source for carbon and nitrogen, through photoautotrophy. A study by Cockell and Stokes determined that the annual primary production of hypoliths on Devon Island (in the Arctic) was approximately similar to that estimated for primary production of plants, lichens and bryophytes. Using the uptake of radio-labelled sodium bicarbonate, this study estimated a productivity of $0.8 \pm 0.3 \text{ g m}^{-2} \text{ y}^{-1}$ in the Arctic, while a previous study estimated productivity of bryophytes at $1.0 \pm 0.4 \text{ g m}^{-2} \text{ y}^{-1}$ on Devon Island (Bliss & Gold, 1994). Assuming an analogous relationship between the Arctic and Antarctic, an assumption may be made that hypoliths play an important role in primary production, although this is subject to further investigation.

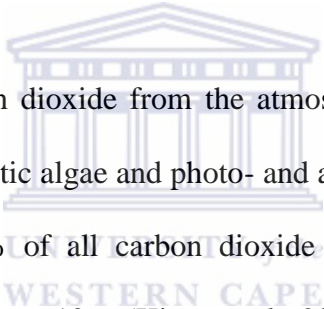
In view that these communities represent significant players in primary productivity, an understanding of the complement of functional genes in the community would shed valuable insight on biogeochemical cycling in Antarctic soil biotopes. It has been recently shown that these communities are homogenous in composition (Pointing *et al.*, 2009). The relative simplicity of hypolithic communities could offer a suitable platform for assessing biogeochemical cycling.

1.3 Global and Regional Carbon Ecology

It is generally believed that there is a continued general increase in the level of anthropogenic carbon dioxide (CO₂) on the earth's atmosphere (Tabita *et al.*, 2007). These levels continue to steadily increase, with a predicted consequence being potential warming of the earth. It has been speculated that living organisms and specific biological macromolecules evolved to catalyze the energy which is required to chemically convert inorganic carbon to reduced organic carbon (Tabita *et al.*, 2008a). Plants, both terrestrial and marine, and specialized microbes developed the ability to remove and assimilate considerably large amounts of CO₂ from the atmosphere and consequently form the necessary organic carbon skeleton required to sustain the biosphere (Tabita *et al.*, 2007).

Average global surface temperatures have increased by 0.8°C since the late nineteenth century, with 11 of the 12 warmest days ever recorded having occurred since 1995 (IPCC, 2007). The average temperature on earth is projected to increase by 1.5 – 5.8°C during the Twenty First Century (IPCC, 2001). Various models projecting climate change have predicted that polar regions will exhibit the most dramatic effects of continental warming (Doran *et al.*, 2002).

Global increases in temperature have been largely attributed to green-house gases (GHGs) such as methane, ozone, nitrous oxide, and carbon dioxide. Anthropogenic activities such as the burning of fossil fuels, and deforestation have been ascribed as additional potential causes of increasing temperatures (Lal, 2008). The levels of CO₂, a principal GHG, have increased in concentration by 12% in the earth's atmosphere, since 1850. This time period coincides with both the industrial revolution and a period of significant increase in population (IPCC, 2007). Although man's assumed role in contributing to global warming may appear significant, it accounts for only 4-5% (2.8 Pg per year out of a 52.8 Pg per year total) of the total biological contributions to atmospheric carbon dioxide (IPCC, 2001; King *et al.*, 2006; United Nations, 2009).



The biological removal of carbon dioxide from the atmosphere occurs principally through carbon fixation by plants, eukaryotic algae and photo- and autotrophic prokaryotes. Microbial decomposition accounts for 86% of all carbon dioxide produced, distantly followed by animal carbon dioxide production at 10% (King *et al.*, 2006). It follows, therefore, that in order to better understand the global carbon cycle, it is important to understand the microbial contribution (Shively *et al.*, 2001).

1.4 Biological Carbon Cycle

An important biogeochemical cycle which is central for primary production in almost all ecosystems is the carbon cycle (Shively *et al.*, 1986). CO₂ fixation is the major biogeochemical process in the carbon cycle (Figure 5). The largest source of organic carbon, by far, is from the lithosphere (Shively *et al.*, 2003). It is thus important to convert atmospheric CO₂ into a usable form (Bliss & Gold, 1994). The resulting C is crucial for primary production in the entire ecosystem (Selesi *et al.*, 2007; Zakhia *et al.*, 2008). This

process is particularly important in harsh Antarctic habitats where soils lack nutrients (Bliss & Gold, 1994).

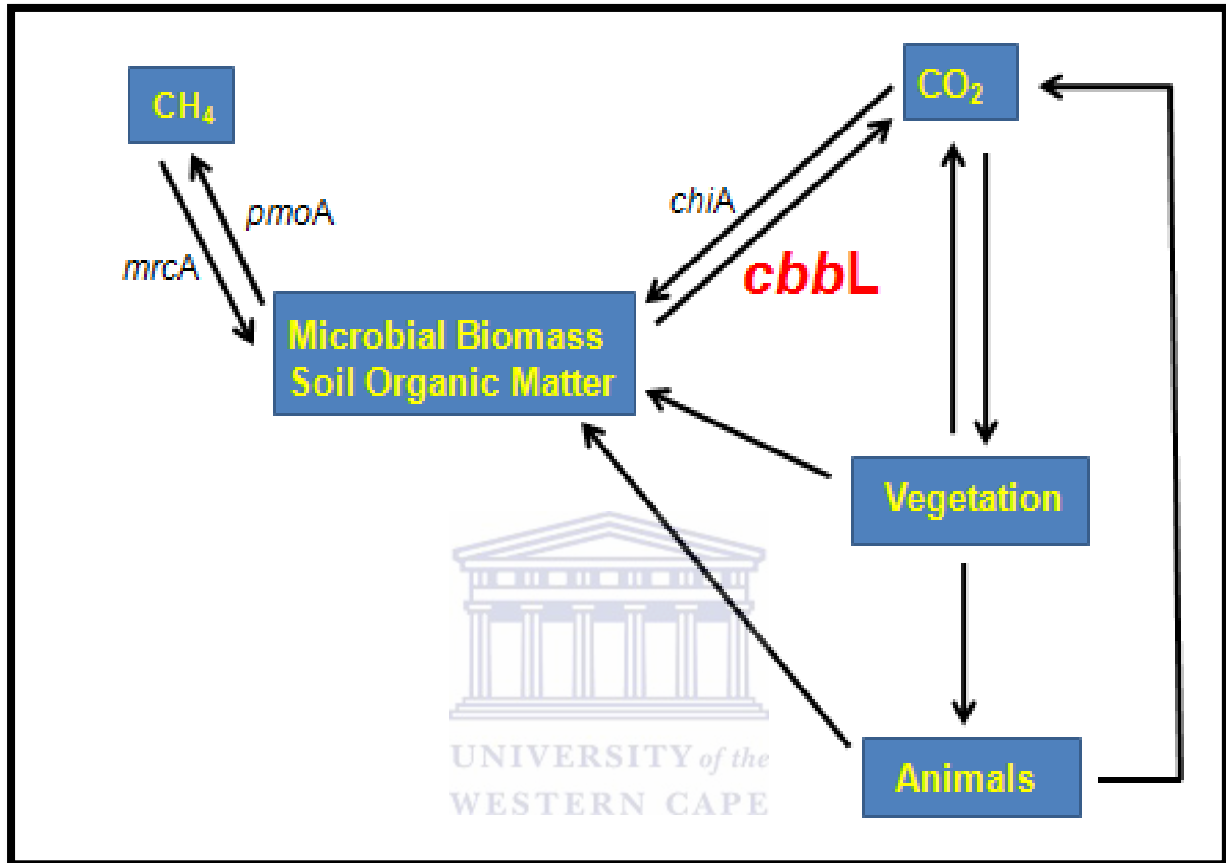


Figure 5: The *cbbL* gene as a marker for CO₂ fixation is involved in the first rate limiting step in the CBB cycle.

In oligotrophic Dry Valley soil, specialized microorganisms are anticipated to play a key role in the assimilation of atmospheric carbon. These specialized microorganisms have the capability to eliminate significant amounts of CO₂ from the atmosphere through different enzymatic schemes (Tabita *et al.*, 2007). There are at least four major metabolic pathways in which microorganisms fix carbon as a source for growth. These are the reductive tricarboxylic acid cycle, the Wood-Ljungdahl acetyl coenzyme A pathway, the hydroxypropionate pathway and the Calvin-Benson-Bassham (CBB) reductive pentose phosphate pathway (Tabita *et al.*, 2007).

1.4.1 The Calvin-Benson-Bassham (CBB) Cycle

The predominant cycle in terrestrial carbon cycling is the Calvin Benson Bassham (CBB) cycle (Figure 6). The CBB cycle involves a series of reactions, which include processes of fixation, reduction and regeneration of the CO₂ acceptor. The resultant reaction results in the enediol form of the sugar bisphosphate ribulose 1, 5-bisphosphate (RuBP) accepting a molecule of CO₂. This reaction is catalyzed by the enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase, catalyzing the primary CO₂ fixation reaction (Tabita *et al.*, 2008b).

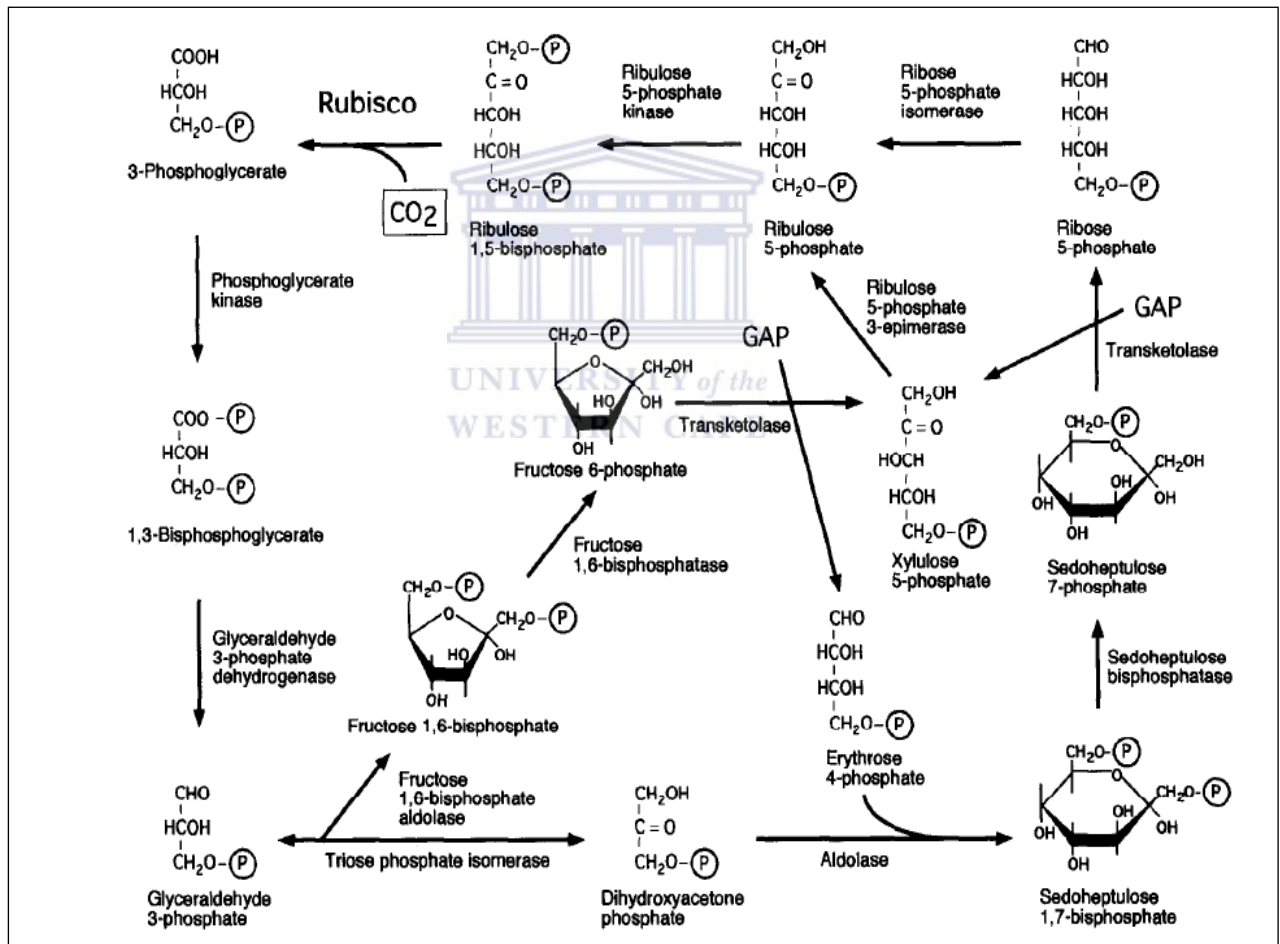


Figure 6: The Calvin-Benson-Bassham cycle (Atomi, 2002)

1.4.2 Ribulose 1, 5-bisphosphate carboxylase/oxygenase (RubisCO)

RubisCO (EC 4.1.1.39) is the most abundant protein on earth and is responsible for catalyzing a crucial step in the fixation of carbon (Tabita *et al.*, 2008a). The ability of RubisCO to catalyze the first rate limiting step for almost all carbon fixation has made it a focus of extensive studies looking at various properties such as the structure, catalytic activity and how it binds CO₂ (Kellogg & Juliano, 1997; Videmšek *et al.*, 2009). RubisCO is found in most autotrophic organisms which include photosynthetic and chemolithotrophic bacteria and archaea, to eukaryotic algae and plants.

RubisCO proteins occur in various forms in nature (summarized in Table 1) (Videmšek *et al.*, 2009). Three of these forms (I, II, and III) are the so called “bona fide” RubisCO proteins and catalyze the same reaction (Figure 7). Another RubisCO, form IV, is actually a homologue and does not catalyze RuBP carboxylation or oxygenation and is therefore called a RubisCO-like protein (RLP) (Tabita *et al.*, 2008a). Form IV RubisCO has similarities (in terms of structure) to the other forms, despite the fact that it is distinguished from them through its lack of catalytic ability.

Table 1: Summary of the properties of different forms of RubisCO (Tabita et al., 2008a)

RubisCO form	Quaternary structure	Types of organisms
I	L ₈ S ₈	plants, algae, proteobacteria, cyanobacteria
II	(L ₂) _n	proteobacteria, dinoflagellates
III	(L ₂) _n	archaea
IV	L ₂	proteobacteria, cyanobacteria, archaea, algae

Form III RubisCO has, as yet, only been found in archaea, many of which are anaerobic extremophiles. RubisCO serves as a means to remove RuBP, produced by isomerization of ribose 1,5-bisphosphate during purine/pyrimidine metabolism (Tabita *et al.*, 2008a). Form II RubisCO has been found in a group of eukaryotes and in various proteobacteria. Composed of only large subunits it is said to only be 30% similar in sequence to Form I large subunits. The ability of Form II RubisCO to discriminate CO₂ from O₂ is less efficient than form I RubisCO. It is frequently found that organisms often contain both Form I and Form II RubisCO. In such cases Form II RubisCO would be the less preferred means of acquiring carbon, but would serve a role of allowing CO₂ to be employed as an electron acceptor balancing the intracellular reduction-oxidation potential when organic carbon is oxidized (Dubbs & Tabita, 1998).

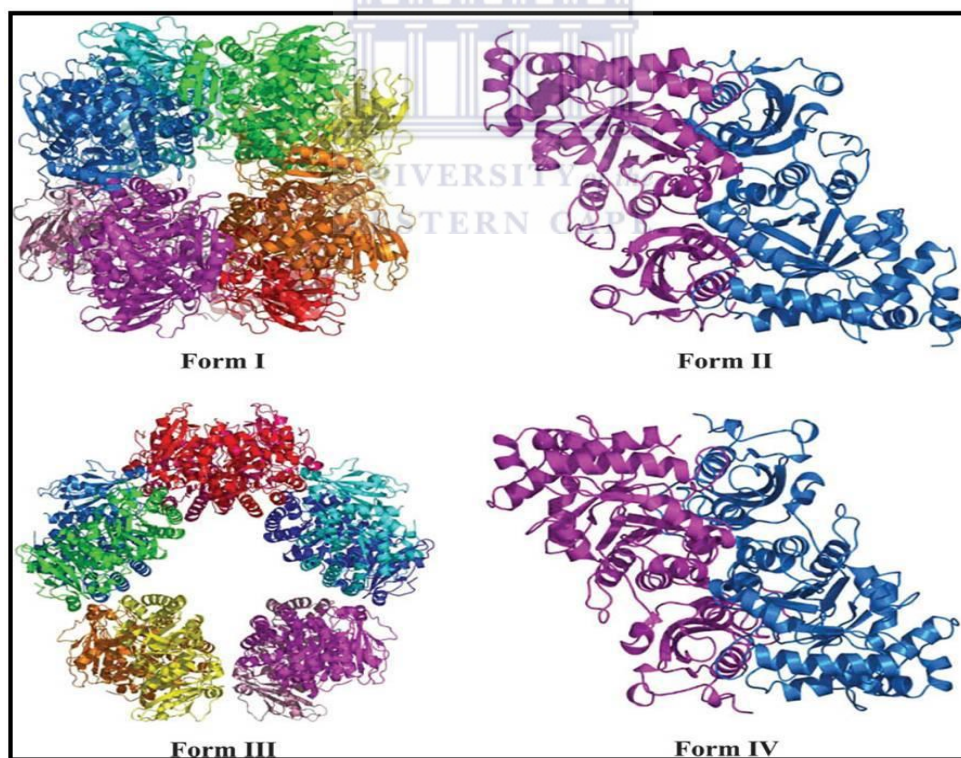


Figure 7: Cartoon representative structures of different Forms of RubisCO (Tabita, 1999)

Form I, is the most abundant and most widely distributed of all the forms of RubisCO (Tabita *et al.*, 2008a). The Form I large subunit is encoded by either the *rbcL* or the *cbbL* genes, with the small subunit being encoded by the *rbcS* or *cbbS* gene (Tabita, 1999). The distinction between *rbc* and *cbb* is made on whether the genes of the operon are controlled by CbbR, whose product controls transcription of the *cbb* operons (Andrew *et al.*, 2005; Dubbs *et al.*, 2000). As such, all the genes carry the prefix *cbb* to denote the fact that they are CBB pathway structural genes (Tabita *et al.*, 1992). RubisCO operons, which do not carry other structural genes, such as cyanobacteria and plant operons, are termed *rbc* (Gibson & Tabita, 1997).

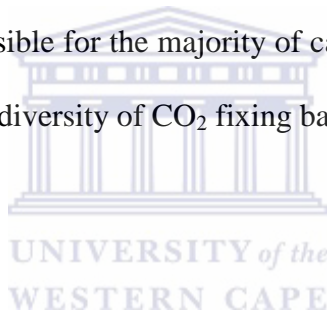
Form I RubisCO may be divided into two main groups, namely the “red” and “green” categories. Broadly, red-like enzymes are found in non-green algae as well as α –and β -proteobacteria, while the green-like enzymes are found in cyanobacteria as well as α - and β -Proteobacteria. In turn, these may be further divided into four subclasses which are termed IA and IB, IC and ID, respectively (Selesi *et al.*, 2005; Tabita, 1999)

The Form I *cbbL* or *rbcL* gene has been used for phylogenetic analyses (Kallersjo *et al.*, 1998; Selesi *et al.*, 2005; Sinigalliano *et al.*, 2003). Use of these genes as probes could potentially facilitate understanding of the key role played by microorganisms in the global and regional carbon ecology, as they are involved in nearly all processes in the carbon cycle (Selesi *et al.*, 2005). Moreover, the role played by soil microorganisms in the ecology of carbon remains poorly understood, more so, within the context of the Arctic and Antarctic, which are the epicenters of ‘global warming’.

1.4.3 Diversity of potential CO₂ fixing bacteria in soil

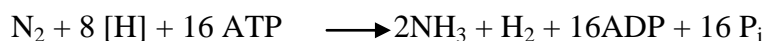
Previous studies have reported an abundance of CO₂ fixing bacteria in soil (Clegg, 1993; Selesi *et al.*, 2005; Sinigalliano *et al.*, 2003; Tolli & King, 2005; Xu & Tabita, 1996) from a wide array of habitats. These studies have been principally based on the RubisCO large subunit gene (*rbcL/cbbL*) and deduced amino acid sequences which are ideally suited as phylogenetic markers).

Soil is an important reservoir for the terrestrial carbon cycle and act as a source or sink for carbon compounds such as CO₂ and CH₄ (Selesi *et al.*, 2005). Soils with their associated uses such as agriculture, are significant contributors to the ‘greenhouse effect’ (Smith *et al.*, 2003). Since microorganisms are responsible for the majority of carbon fixation in soils, it would be beneficial to fully understand the diversity of CO₂ fixing bacteria in these habitats.



1.5 Nitrogen Fixation

The nitrogen cycle is another important biogeochemical cycle (Figure 8). Biological nitrogen fixation (BNF) is a major biogeochemical process resulting in the catalysis of atmospheric nitrogen into ammonium (equation below). The resultant ammonium is biologically available and forms the structural basis of proteins, nucleic acids, and many other biomolecules (Seefeldt *et al.*, 2009).



A minority of N₂ required for human use is provided through the industrial Haber-Bosch process, with the majority of N₂ fixation resulting from fixation by microorganisms (Seefeldt

et al., 2009). The total amount of biologically fixed nitrogen exceeds 2×10^{13} g/year (Falkowski, 1997).

The equilibrium resulting from N_2 fixation and the reverse process of denitrification, can determine the net biologically available N for the biosphere (Zehr *et al.*, 2003). In such a way, N_2 interacts with other biogeochemical cycles to control the N status of the ecosystem (Zehr *et al.*, 2003).

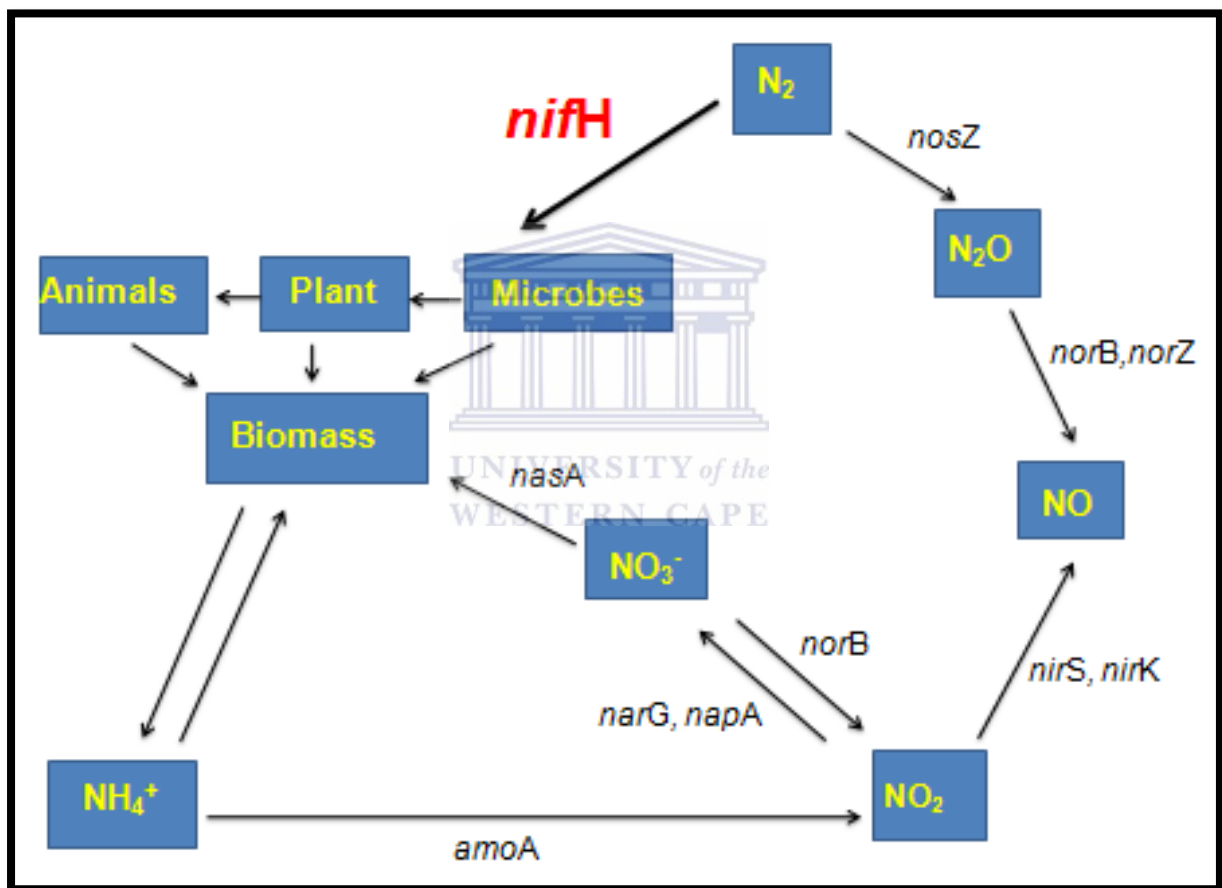
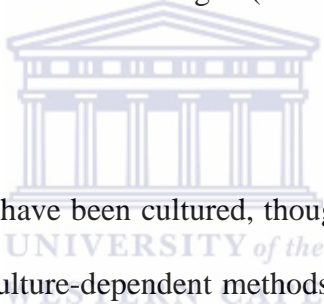


Figure 8: Nitrogen cycling and genes involved in biogeochemical N cycling. The *nifH* gene represents the rate limiting step in the nitrogen cycle

Fixation of atmospheric nitrogen is carried out by a limited, phylogenetically diverse group of proteobacteria, green sulphur bacteria, firmicutes, spirochaetes, archaea and cyanobacteria (Zhang *et al.*, 2008).

1.5.1 Diazotrophs

Diazotrophs are defined as bacteria capable of fixing atmospheric nitrogen (Postgate, 1998). Diazotrophs are not limited to a specific bacterial taxonomic group, and within a single genus, both N-fixation and non-N-fixing species may occur (Postgate, 1998). In natural ecosystems biological nitrogen fixation (BNF) by free-living, associated, and symbiotic diazotrophs is the most important source of nitrogen (Berman-Frank *et al.*, 2003; Russelle & Birr, 2004).

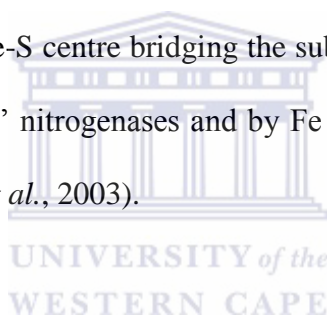


Diazotrophs from various groups have been cultured, though growth conditions places some constraints on the usefulness of culture-dependent methods to easily and accurately evaluate diversity (Postgate, 1998). It is also likely that certain diazotrophs are unculturable; making culture independent approaches a more attractive option to assess diversity (Jenkins *et al.*, 2004; Steward *et al.*, 2004). In order to assess the diversity of diazotrophic microorganisms, dinitrogen fixation genes have previously been used as functional markers (Zehr *et al.*, 2003).

All dinitrogen reductase genes *nifH*, *vnfH* and *anfH* are highly conserved in structure. The *nifH* gene has predominantly been the focus of culture-independent approaches to assess diazotroph diversity (Zehr *et al.*, 1997; Zehr & Ward, 2002).

1.5.2 Nitrogenase

Nitrogenase is the primary enzyme responsible for diazotrophy (Postgate, 1998). The heterotetramer metalloprotein dinitrogenase $\alpha_2\beta_2$ and the homodimer metalloprotein dinitrogenase reductase γ_2 are the two proteins which form nitrogenase. The dinitrogenase (Mo-Fe protein) catalyses the reduction of N_2 , and the dinitrogenase reductase (Fe protein) donates electrons of very low redox potential to dinitrogenase (Luque & Forchhammer, 2008). The homodimer (Fe-protein) is composed of a single Fe_4S_4 cluster bound between identical subunits which are approximately 32- 40 kDa (Berman-Frank *et al.*, 2003). The heterotetramer (Mo-Fe- protein) is approximately 25 kDa and contains a P cluster and the M centre (FeMo co-center) (Berman-Frank *et al.*, 2003). The so-called ‘conventional’ nitrogenases contain Mo in the Fe-S centre bridging the subunits. The metal is replaced by V (vanadium) (*vnfH*) in ‘alternative’ nitrogenases and by Fe (*anfH*) in the ‘second alternative’ nitrogenases (Thiel, 1993; Zehr *et al.*, 2003).



The conserved nature of *nifH* provides a convenient basis for developing probes in order to study nitrogen fixation genes in diverse organisms (Zehr *et al.*, 2003). Sequencing data suggest that broad taxonomic groups including closely related species can potentially be distinguished by comparison of *nifH* DNA and deduced amino acid sequences (Kirshtein *et al.*, 1991; Zehr & McReynolds, 1989).

An assumption is made that the presence of *nif* implies N_2 fixing capability, because the multiple genes involved in regulation and assembly occupy a large region of the genome (Zehr *et al.*, 2003). As previously observed, nitrogenase gene expression is highly regulated from transcription to post translational protein modification. Transcription of the *nifHDK*

operon is an appropriate marker for N₂-fixing conditions because it is not constitutively expressed and is regulated in response to factors that control N₂ fixation (Zehr *et al.*, 2003).

1.5.2.1 Dinitrogenase and Dinitrogenase Reductase

Nitrogenase consists of two multi subunit proteins: dinitrogenase reductase, encoded by the *nifH* gene, and dinitrogenase, encoded by the *nifD* and *nifK* genes. Studies on the diversity of *nifH*-containing plankton in the open ocean have revealed a diverse suite of potential diazotrophs, including cyanobacteria and proteobacteria, as well as *nifH* sequences from anaerobic bacteria (Church *et al.*, 2005). Dinitrogenase reductase is encoded by the *nifH* gene and dinitrogenase by the *nifD* (α) and *nifK* (β) genes (Rees *et al.*, 2005; Zehr *et al.*, 1997).



1.5.3 Major *nifH* gene clusters

Nitrogenase genes are clustered into four major groups, designated as Clusters I-IV (Table 2) (Chien & Zinder, 1996).

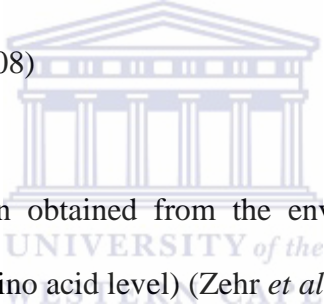
Table 2: Composition of major nitrogenase gene clusters (Zehr *et al.*, 2003)

Cluster	Composition
I	'conventional' Mo containing <i>nifH</i> and some <i>vnfH</i>
II	'second alternative' non-Mo, non-V containing <i>anfH</i> including nitrogenases from certain Archaea.
III	<i>nifH</i> sequences from a diverse group of distantly related strict anaerobes such as clostridia
IV	divergent loosely coherent group of <i>nif</i> -like sequences from Archaea and distantly related chlorophyllide reductase genes

Most microorganisms, capable of diazotrophy, have multiple copies of the nitrogenase genes or homologues of nitrogenase genes (Cluster IV). For instance *Clostridium pasteurianum* has previously been shown to have a *nifH* gene family sequences in Cluster II and III (Wang *et al.*, 1988). With some cyanobacteria also having multiple copies of nitrogenase, including a *vnfH* as well as a second distinct *nifH* (Thiel, 1993). In spite of these exceptions, the *nifH* gene remains the best functional marker for nitrogen fixation.

1.5.4 Diversity of potential nitrogen fixing microorganisms

Although a number of studies have been based on cultivation of diazotrophs, the majority of *nifH* sequences deposited in databases are from uncultured organisms (Simonet *et al.*, 1991; Zehr *et al.*, 2001; Zhang *et al.*, 2008)



nifH sequences which have been obtained from the environment have been remarkably similar (> 94% identical at the amino acid level) (Zehr *et al.*, 2003). In 1997, it was noted that cyanobacterial *nifH* genes form clusters (Zehr *et al.*, 1997). Trees which have been constructed for *nifH* and 16S rRNA genes based on the same lineages represented by both genes have generated highly consistent trees (Zehr *et al.*, 2003). This indicates a strong evolutionary congruency between both genes, and therefore suggests that the *nifH* gene may be validly used as a phylogenetic marker.

In the Antarctic the microbial composition of terrestrial ecosystems has been the subject of intensive recent study (Babalola *et al.*, 2009; Rochelle *et al.*, 2009; Zeglin *et al.*, 2009). However, little is known about the functional interactions among organisms and the environment (Nakatsubo & Ino, 1986), especially in the context of carbon and nitrogen fixation.

1.6 Aims and Objectives

In understanding the complexity of the carbon and nitrogen cycles it is important to appreciate the close interplay in terrestrial systems (Figure 9). For instance, nitrogen has previously been reported to be a potentially limiting factor in carbon assimilation (Falkowski, 1997; Nakatsubo & Ino, 1986).

Antarctica is uniquely suited for investigating the close relationship between these cycles (C and N). The relative simplicity offered by its various biotopes, and the ability of microorganisms to respond speedily to environmental changes, enable probing of these relationships (Melody *et al.*, 2001). It is predicted that climate change may lead to an increase in the colonization of new species to Antarctica (Convey, 1996). This is expected to lead to increased diversity, biomass, trophic complexity and habitat structure. The ensuing competition may possibly result in a loss of current Antarctic species and communities (Barrett *et al.*, 2008; Convey, 1996).

This study presents, for the first time, a ‘snap-shot’ overview of key anabolic genes (*nifH* and *cbbL*) implicated in C and N cycling processes. Through the use of culture-independent techniques, community compositions are probed using the *nifH* and *cbbL* genes as functional markers.

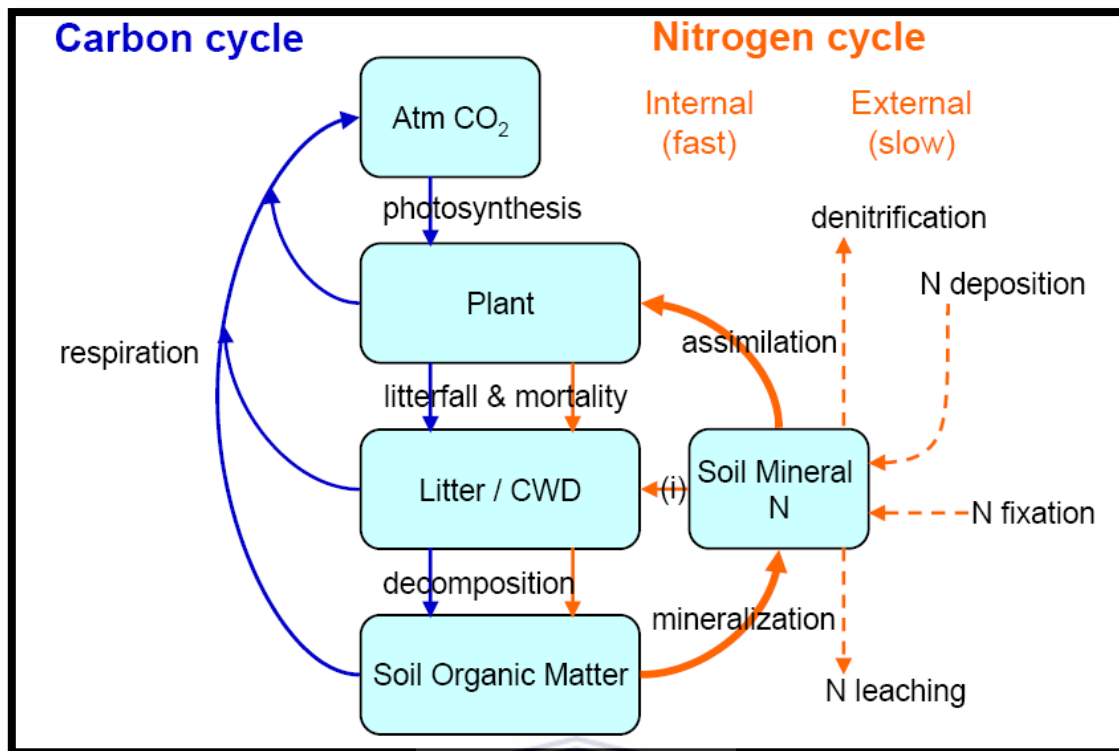


Figure 9: Feedback pathways Interaction between C and N cycle. There is a close relationship between mineral nitrogen and atmospheric CO₂ fixation through the action of plants. In Antarctic environments microorganisms are expected to fulfill the role normally taken by plants (Thornton *et al.*, 2009).

1.6.1 Aims

To investigate the diversity of key anabolic genes in Antarctic hypolithic communities

1.6.2 Objectives

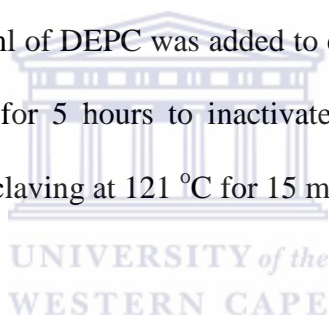
- To obtain a snapshot of the total bacterial diversity through 16S rRNA gene DGGE analysis.
- To determine the diversity and expression of *nifH* genes in a hypolithic community (nitrogen fixation).
- To determine the diversity *cbbL* (both red-like and green-like) genes in a hypolithic community (carbon fixation).

Chapter 2: Materials and Methods

This chapter details experimental protocols/reagents used during the course of this study. All reagents used were of the highest molecular biology grade and were purchased from vendors employing standard ISO protocols to ensure quality control (Appendix 1). Unless otherwise indicated all reagents were purchased from Sigma.

2.1 General procedures for the preparation of reagents and glassware for RNA work

All water or salt solutions (not containing Tris) used in RNA preparation were treated with diethylpyrocarbonate (DEPC). 1ml of DEPC was added to each litre of solution to be treated. Glassware was baked at 140°C for 5 hours to inactivate RNase. All media, buffers, and glassware were sterilized by autoclaving at 121 °C for 15 minutes.



2.2 Media

The following media were used;

i) **Luria-Bertai agar (LB) ((Sambrook & Russell, 2001): values in g/L)**

Yeast extract	10.0 g
Tryptone	5.0 g
NaCl	10.0 g
Agar	15.0 g

The pH was adjusted to 7 and before autoclaving and the medium was supplemented with 100 mg/ml of ampicillin and 100 mM MgCl₂ after cooling to less than 55 °C.

ii) LB agar Medium (Luria-Bertani Media) ((Sambrook & Russell, 2001):

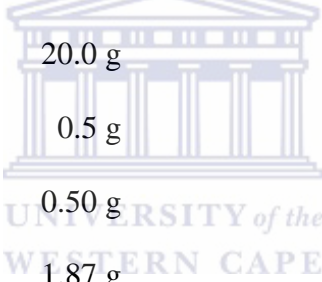
values in g/L)

Tryptone	10.0 g
Yeast extract	5.0 g
NaCl	10.0 g
Agar	15.0 g

The pH was adjusted to 7 with 5 N NaOH, made up to 1L.

iii) SOB agar (LB) ((Sambrook & Russell, 2001): values in g/L)

Tryptone	20.0 g
Yeast extract	0.5 g
NaCl	0.50 g
250 mM KCl	1.87 g
Agar	15.0 g



The pH was adjusted to 7 before autoclaving and the medium was supplemented with 100 mg/ml of ampicillin and 100 mM MgCl₂ after cooling to less than 55 °C.

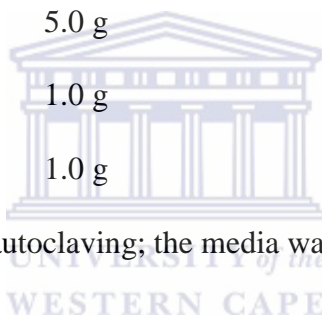
iv) SOC Media ((Sambrook & Russell, 2001): values in g/L)

Tryptone	20.0 g
Yeast extract	5.0 g
NaCl	0.5 g
250 mM KCl	10.0 ml

The pH was adjusted to 7 before autoclaving; the medium was cooled to ~ 50 °C and the following filter sterilized and added aseptically, 5 ml of 2 M MgCl₂ and 20 ml of 1M glucose; the media was made up to 1L.

v) TGY Media ((Sambrook & Russell, 2001): values in g/L)

Tryptone	5.0 g
Glucose	1.0 g
K ₂ HPO ₄	1.0 g



The pH was adjusted to 7 before autoclaving; the media was made up to 1L.

vi) ATCC medium ((Sambrook & Russell, 2001): values in g/L)

KH ₂ PO ₄	0.2 g
K ₂ HPO ₄	0.8 g
MgSO ₄ .7H ₂ O	0.2 g
CaSO ₄ .2H ₂ O	0.1 g
FeCl ₃	2.0 mg
Na ₂ MoO ₄ .2H ₂ O	Trace
Yeast extract	0.5 g
Sucrose	20.0 g

The pH was adjusted to 7 before autoclaving; the media was made up to 1L.

2.3 Environmental Sample collection

Samples were collected in sterile plastic bags by Prof. Don Cowan from Antarctic hypolithons during December to January of 2008/9. Samples meant for RNA extractions were stored in RNA*later*TM. Samples were immediately placed on dry-ice and stored at -80°C until further use.

2.4 DNA/RNA Extractions

2.4.1 Modified Miller Protocol

In order to extract DNA from hypolithic soils a protocol first described by Miller and colleagues (Miller *et al.*, 1999) was modified as follows. Aliquots of soil samples (0.5 g) were thawed and mixed with 0.5 g of quartz sand in a 2 ml screw cap (Sigma S-9887). 300 µl of sodium phosphate (100 mM, pH 8.0) and lysis buffer (100 mM NaCl, 500 mM Tris [pH 8.0], 100 g/l SDS) were added, and the tube was inverted to mix. After the addition of 300 µl of chloroform/isoamyl alcohol (24/1, v/v) the mixture was vortexed at maximum level for 2 min, and the cell debris was pelleted by centrifugation (15,000 x g, 5 min). To the supernatant ammonium acetate (7M) was added to a final concentration of 2.5 M. The tube was subsequently shaken by hand to mix, followed by centrifugation (15,000 x g, 10 min). The DNA was precipitated from the supernatant by addition of 0.6 volumes of isopropanol, incubation at room temperature for 15 min. The DNA was pelleted by centrifugation (15,000 x g, 10 min) and washed with 1 ml of 70% ethanol. The DNA was then air dried and dissolved in 50 µl of TE (10 mM Tris, 1 mM EDTA, pH 8).

2.4.2 DNA extraction from bacterial cultures

Bacterial culture was grown over night in liquid media. Bacterial suspension was centrifuged for 5 min at 4500 x g to pellet the bacteria. The pellet was re-suspended in 1 ml of extraction buffer (50 mM NaCl, 50 mM Tris-HCl, pH 7.6, 50 mM EDTA and 5% SDS [pH 8.0]) and transferred to a 2 ml microcentrifuge tube containing 0.4-0.5 g of glass beads (0.1 mm diameter) and vortexed for 2 min. 300 µl of phenol and chloroform/isoamyl alcohol (25:24:1, v/v) were added before vortexing for a further 20 seconds. After centrifugation at 14,000 x g for 3 min the upper aqueous phase was transferred to a new 2 ml tube and re-extracted with 500 µl of chloroform. DNA was precipitated with 0.1 volume of 3M sodium acetate solution and 0.7 volumes of isopropanol. DNA was precipitated by centrifugation at 14,000 x g for 30 min at 10°C. The supernatant was discarded by aspirating the isopropanol. The pellet was washed with 70% ethanol and the DNA was re-pelleted by centrifugation for 5 min. After removing the ethanol the pellet was allowed to air dry at room temperature for 5 min and eluted in TE (10 mM Tris, 1 mM EDTA, pH).

2.4.3 RNA extraction using MoBio Soil RNA kit

The protocol was followed as specified by manufacturer with minor modifications. To a 15 ml Bead Tube, 2 g of soil, 2.5 ml of Bead Solution and 0.25 ml of Solution SR1 was added before vortexing the solution to mix. 0.8 ml of Solution SR2 was added before vortexing for a further 5 min at maximum speed. 3.5 ml of phenol: chloroform: isoamyl alcohol (25:24:1, pH8.0) was then added before vortexing until the disappearance of the biphasic layer. The 15 ml Bead Tube was then vortexed at maximum speed for a further 10 min. Centrifugation at 6000 x g for 10 min (at 4°C) was applied. The resulting supernatant was carefully transferred to a new bead tube. 1.5 ml of Solution SR3 was added to the 15 ml Bead Tube containing aqueous phase prior to vortexing at maximum speed. After incubating the Bead Tube at 4°C

for 10 min, the mixture was centrifuged at 6,000 x g for 10 min (at 4°C). The supernatant was then transferred to a new 15 ml Bead Tube with 5 ml of Solution SR4 being added before incubation at -20°C for 30 minutes. Centrifugation at 6000 x g for 30 min was applied at 10°C. The supernatant was decanted and the 15 ml Bead Tube was inverted and placed on a paper towel for 5 min. 1 ml of Solution SR5 was added before being applied to the provided RNA Capture Column. The column was then washed with Solution SR5. Solution SR6 was applied to the column and collected in a new tube. Eluted RNA was transferred to a 2.2 ml collection tube and 1 ml of Solution SR4 was added before centrifuging the mix for 15 min at 13,000 x g at room temperature to pellet the RNA. The supernatant was decanted and the 2.2 ml collection tube was air dried. The RNA pellet was re-suspended in 50 µl of Solution SR7.



UNIVERSITY of the
WESTERN CAPE

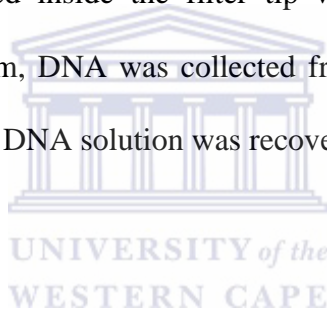
2.4.4 Modified DNA/RNA protocol using CTAB lysis buffer

The protocol was modified by Katherine Koyné from one previously described Griffith and colleagues (Griffiths *et al.*, 2000). Soil samples stored in RNAlater® were thawed, and mixed well. 1 ml of soils was then added to 2 ml screw-cap tube and centrifuged at 4000 rpm, 3 min at room temperature. The RNAlater® was removed before mixing soil with 1 ml of DEPC-treated PBS. Centrifugation was performed as above before removing PBS. 0.8 ml of CTAB Buffer (100 mM Tris-HCl [pH8.0], 1.4 M NaCl 20 mM EDTA [pH8.0], 2 % [w/v] CTAB, 1 % [w/v] polyvinylpyrrolidone) was added to ice-cold beads (0.5 g each of 0.1 mm glass and 0.5 mm zirconia/silica) and 0.5 g of wet soil. The mixture was vortexed before adding 3 µl of β-mercaptoethanol. Vortexing was done in 3 X 30 sec pulses. 0.8 ml of chloroform:isoamylalcohol (24:1) was added before vortexing and rocking at room temperature for 15 min. Centrifugation (15 min, 4°C) was applied at 15,000 rpm before transferring the supernatant to a new tube and adding 1 volume of chloroform: isoamyl

alcohol (24:1) and vortexing well. The mixture was centrifuged again as previously described, before adding 0.5 volume of NaCl and 1 volume isopropanol. The solution was incubated at -80°C overnight. Nucleic acids were pelleted by centrifugation (15,000 rpm for 15 min at 4°C). The resulting pellet was washed with 70% ethanol, and dried on a paper towel. Nucleic acids were resuspended in 20 µl of RNase free water.

2.4.4 Purification using Polyvinylpolyrrrolodone (PVPP) minicolumns

PVPP spin mini-columns were assembled using a P20 tip (end of tip cut) placed inside a 0.6 ml PCR tube (with end and lid cut) enclosed in a 1.5 ml tube (lid detached). A solution of 50% PVPP in TE was deposited inside the filter tip with 50 µl DNA. Columns were centrifuged for 20 min at 3.3 rpm, DNA was collected from the base of the tubes and the process was repeated until a clear DNA solution was recovered (typically twice).



2.5 Bacterial Strains

Two Proteobacteria strains were sourced from the United States Department of Agriculture's Agricultural Research Service Culture Collection. *Azotobacter vinelandii* (B-14641) and *Azospirillum brasilense* (B-14647) were grown as instructed.

2.6 Polymerase Chain Reaction

PCR amplifications were performed in 0.2 ml thin-walled tubes using a thermocycler equipped with a heated lid (e.g., PCR Sprint Temperature Cycling System, Thermo Hybaid, Ashford, GB). A standard 50 μ l reaction contained 1 X PCR buffer (10 X being 200 mM Tris pH 8.8, 100 mM KCl, 100 mM $(\text{NH}_4)_2\text{SO}_4$, 20 mM MgSO_4 , 1% (w/v) Triton X-100), 0.2 mM each of dATP, dCTP, dGTP and dTTP, 0.5 μ M of each primer, an appropriate amount of *Taq* DNA polymerase, and 25 ng of metagenomic DNA as template. For DNA amplifications from whole cells, small amounts of freshly grown colonies were transferred to the reaction tubes using a 10 μ l pipette. The primers used as well as the cycling conditions applied are given in Table 3.



Table 3: Primer combinations, cycling conditions and specificity of primers used in this study.

Primer Set	Sequence (5' to 3')	Amplification Cycle	Specificity	Reference
341FGC 534R	CCTACGGGAGGCAGCAG ATTACCGCGGCTGCTGG	94°C/4 min 20x (94°C/45 s - 65°C*/45 s - 72°C/60 s) #	Most bacteria	(Muyzer <i>et al.</i> , 1993)
PolF PolR	TGC GAY CCS AAR GCB GAC TC ATS GCC ATC ATY TCR CCG GA	98°C/30 sec 30x (98°C/10 s - 67°C/15 s - 72°C/45 s)*	<i>nifH</i> gene	(Poly <i>et al.</i> , 2001)
cbbLRF cbbLRR	AAG GAY GAC GAG AAC ATC TCG GTC GGS GTG TAG TTG AA	94°C/4 min 30x (94°C/45 s - 60°C/45 s - 72°C/60 s)#	<i>cbbL</i> gene "red like" form	(Selesi <i>et al.</i> , 2005)
cbbLGF cbbLGR	GGC AAC GTG TTC GGS TTC AA TTG ATC TCT TTC CAC GTT TCC	94°C/4 min 20x (94°C/45 s - 65°C*/45 s - 72°C/60 s)#	<i>cbbL</i> gene "green-like" form	(Selesi <i>et al.</i> , 2005)
M13 For M13 Rev	CCCAGTCACGACGTTGTAAAACG AGCGGATAACAATTTACACAGG	94°C/10 min 35x (94°C/30 s - 64°C/30 s - 72°C/[var]) #	Cloning vector pGEM [®] T Easy	
pJET1.2 forward pJET1.2 reverse	CGACTCACTATAGGGAGAGCGGC AAGAACATCGATTTTCCATGGCAG	94°C/10 min 35x (94°C/30 s - 64°C/30 s - 72°C/[var]) #	Clone Jet PCR Cloning Kit	pJET Cloning Kit

* Phusion *Taq* polymerase used # IMBM lab *Taq*

2.7 Analytical procedures

Spectrophotometry and agarose gel electrophoresis were used to quantify and verify both DNA and RNA in this study. Reference DNA fragments were developed from PCR amplification of either known 16S rRNA genes or from the digestion of DNA with *Pst*I restriction enzyme. In some cases commercial markers from Fermentas were used.

2.7.1 Agarose gel electrophoresis

Nucleic acid fragments were separated by agarose gel electrophoresis. Total genomic DNA fragments and PCR amplicons were separated in 1% - 2.5% agarose gels, prepared in 0.5 X TAE buffer (Sambrook & Russell, 2001). Ethidium bromide (0.5 µg/ml), was added to the agarose gels during preparation for the staining and visualization. Samples were prepared by mixing with 6x concentrated loading buffer (20% (v/v) glycerol and 5 mg/ml bromophenol). Electrophoresis was performed in 0.5x TAE buffer at 100 V. DNA bands were sized according to their migration in the gel as compared to DNA molecular weight markers (e.g., DNA cut with *Pst*I restriction enzyme). Gels were visualized via ultraviolet (UV) light illumination and photographed with a digital imaging system (Alphamager 2000, Alpha Innotech, San Leandro, CA).

2.7.2 Fluorometry (Qubit™)

The fluorometer was used to measure the concentration of obtained DNA/RNA extractions following the protocol as outlined by the manufacturer. All reagents to be used were incubated at room temperature before use. Preparation of assay tubes was

carried out as indicated in Table 4. A Quanti-iT™ Working Solution was set up by diluting the Quanti-iT™ reagent 1:200 in Quanti-iT™ buffer. 200ul of Working Solution was set-up for each sample and standard.

Table 4: Preparation of assay tubes was carried out as follows

	Standard	User Sample
Volume of Working Solution	190 µl	180-199 µl
Volume of Standard to add	10 µl	-
Volume of DNA to add	-	1-20 µl
Total Volume in each Assay tube	200 µl	200 µl

All tubes were vortexed for 2-3 sec and incubated at room temperature for 2 min. The tubes were then read in a Qubit™ fluorometer.

2.7.3 Spectrophotometry

The DNA/RNA concentrations (calculated as OD_{260 nm} x 50 ng/µl) and purity (ratio OD_{260 nm}/ OD_{280 nm}) were measured using the NanoDrop® ND-1000 UV-Vis Spectrophotometer (NanoDrop Technologies, USA). The solvent used for DNA elution was the reference sample for the determination of the DNA spectro-photometric parameters.

2.8 Gel Extraction and Purification

PCR fragments were separated on TAE agarose gels by electrophoresis in 0.8% as described above (section 2.5.1). Using a sterile scalpel, bands were excised from the gel and placed in 1.5 ml tubes. Gel extraction and purification was performed as detailed by the manufacturer.

2.9 Denaturing Gradient Gel Electrophoresis (DGGE)

DGGE plates were cleaned thoroughly with methanol then twice with ethanol to remove all traces of grease before assembly. 20 µl of PCR-DGGE products (approximately 500 ng DNA) were separated by DGGE as described by Muyzer *and* colleagues (Muyzer *et al.*, 1993). Urea-formamide (Fluka) gel denaturing gradients were developed using the Bio-Rad Gradient-former (Bio-Rad, Hercules, USA). 0.5% (w/v) APS and 0.02% (v/v) TEMED were added to the acrylamide: bisacrylamide (37.5:1(w/w)) solutions, as catalysts for gel polymerization prior to gradient development. Universal bacterial DNA was separated on 30-70% urea-formamide gradients (A 100% urea-formamide solution contains 7M urea and 40% (v/v) formamide). DGGE was performed using the Bio-Rad Electrophoresis apparatus on 16.5mm x 16.5mm x 1mm thick 9% (w/v) polyacrylamide gels, at 100V for and 60 °C for 16 h in 1 X TAE buffer). After electrophoresis, the gels were stained using 0.5µg/ml ethidium bromide in 1XTAE for 15 min and destained in 1 X TAE for 30 min before visualizing and capturing the image using the Alphaimager 3400 Imaging System UV transilluminator (AlphaInnotech Corporation™ San Leandro, CA).

2.10 Construction of Clone libraries

2.10.1 Preparation of *Escherichia coli* (*E. coli*) competent cells

Electrocompetent GeneHogg *E. coli* cells were prepared by inoculating a single freshly streaked colony in 20 ml SOB media and cultured overnight at 37⁰C with agitation at 250 rpm. 2 ml of the overnight culture was inoculated into a 250 ml sterile flask and cells were grown at 37⁰C with shaking (250 rpm) for 3.5 - 4 h to an OD_{600nm} of 0.6. Cells were kept on ice and harvested by centrifugation at 4000g for 10 min at 4⁰C. The cells were then re-suspend in a volume of ice-cold sterile water equal to the original culture volume, and then harvested as before. The supernatant was discarded and the cells were re-suspended in ice-cold sterile 10% glycerol and then centrifuged for 15 min at 4000g. After decanting the supernatant, cells were re-suspended in ice-cold sterile 15% glycerol, 2% sorbitol using a volume of 2ml per L initial culture. Cells were harvested by centrifugation at 4000g for 10 min at 4⁰C and then re-suspended in a volume of ice-cold sterile 15% glycerol, 2% sorbitol equal to that of the cell pellet. 50µl volumes of cells were then aliquoted into clean 0.5ml eppendorf tubes. Ethanol at -80⁰C was used to snap freeze the cells. Cells were stored at -80⁰C.

2.10.2 Ligation of PCR products

Ligation reactions were performed using the following kits;

(i) Promega pGEM-T™ Easy

Ligations were carried out using p-GEM T Easy according to manufacturer's instructions. Ligations were carried out in 10 µl volumes. To each tube 5 µl of Rapid ligation buffer was mixed with 1 µl of pGEM-T™ Easy vector. To each experiment tube 3 µl of PCR product was added to the reaction mixture with 1 µl of ligase being added.

(ii) Clone Jet PCR Cloning

For blunt-end PCR products generated from proofreading DNA polymerases (Phusion™ *Taq*), ligation was performed using Clone Jet PCR Cloning Kit. A ligation was set up as per manufacturer's instructions. Ligations contained 10 µl of 2X reaction buffer. 2 µl of PCR products were used in a 3:1 ratio with pJET 1.2 blunt cloning vectors (50 ng/µl). A lethal gene, incorporated into the pJET vector, is disrupted by ligation of a DNA insert into the cloning site. As a result, only cells with recombinant plasmids can propagate thereby eliminating the need for blue/white selection.

2.10.3 Transformation of competent *E. coli* Gene Hogg strain

The electro-competent Gen Hogg *E.coli* cells were transformed with purified DNA. An Eppendorf tube containing 50 µl of electrocompetent cells was removed from -80°C storage and allowed to thaw on ice. 2µl of the ligation mixture (section 2.10.2) was added to the thawed cells and gently mixed. The mixture was incubated on ice for ~ 1

min then pipetted into a pre-chilled 0.1 cm sterile electroporation cuvette (Bio-Rad). Electroporation was performed under the following conditions: 1.8 kV, 25 μ F, 200 Ω on the BioRad Gene Pulser machine. Immediately following electroporation, 1 ml SOB broth was added to the cuvette, the cells transferred to a 15 ml Falcon tube and incubated at 37°C for 1 h with agitation. 100 μ l of cells were plated onto LB-agar plates supplemented with ampicillin (100 μ g/ml), IPTG (20 μ g/ml), and X-Gal (30 μ g/ml). Recombinant transformants were selected by blue/white colour selection based on insertional inactivation of the *lacZ* gene for transformations done using pGEM T-Easy™.

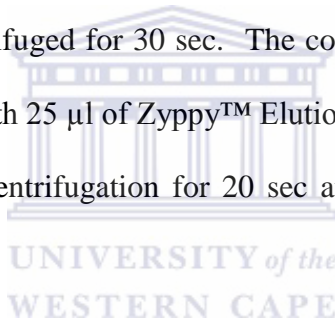
2.11 Screening of clone libraries

2.11.1 Colony PCR

Blue colonies were picked from overnight culture plates using sterile toothpicks and swirled in 50 μ l TE buffer. 2 μ l was directly used as a template for PCR. PCR reaction consisted of 5 μ l of 10X buffer, 4 μ l of 25 mM MgCl₂, 5 μ l of 5 μ M of each primer either M13 or pJET (Table 3), 5 μ l of 1mM dNTP's, 0.5 μ l of *Taq* polymerase (IMBM lab *Taq*) and 2 μ l of the re-suspended colony in a final volume of 50 μ l. Amplification was in an automated thermal cycler (Thermo Hybaid system) with the cycling parameters performed specific for each primer combination as detailed in Table 3. The products were separated by agarose gel electrophoresis (section 2.5.1) and cleaned using the method described in section 2.6 before sequencing.

2.11.2 Plasmid DNA extraction

Plasmid extraction was performed using the Zippy™ Plasmid Miniprep Kit, using manufacturer's instruction with minor modifications. 600 µl of bacterial culture was added to 1.5 ml microcentrifuge tubes. 100 µl of 7X Lysis Buffer was added and mixed by inverting the tube 5 times. 350 µl of cold neutralization buffer was added and mixed thoroughly within 2 min. Centrifugation was applied at 13,000 x g for 4 minutes before transferring supernatant to the Zymo-Spin™ II column. The column was placed in the Collection Tube and centrifuged for 20 seconds. 200 µl of Endo-Wash Buffer was added to the column and centrifuged for 20 sec. 400 µl of Zippy™ Wash Buffer was then added to the column and centrifuged for 30 sec. The column was transferred to a clean 1.5 ml microcentrifuge tube with 25 µl of Zippy™ Elution Buffer being added directly to the column matrix. A final centrifugation for 20 sec at 11,000 X g followed to elute plasmid DNA.



2.11.3 Restriction Fragment Length Polymorphism (RFLP)

2.11.3.1 Analysis of RFLP patterns

RFLP analysis of PCR amplicons was done in 20 µl reaction mixtures containing 100 ng of DNA, 1 U of restriction endonuclease (Fermentas) and the appropriate 10x buffer. The reactions were incubated overnight at 37°C, then the restriction digests were separated using agarose gel electrophoresis (2.5% gels). The appropriate restriction enzymes were selected using WatCut, an on-line tool available on <http://watcut.uwaterloo.ca/>.

2.12 Sequencing and Sequence analysis

2.12.1 Sequencing reactions

Sequencing of plasmid DNA was carried out at the University of Stellenbosch. Sequencing reactions were performed using either M13F or pJETF oligonucleotide primers, for pGEM T Easy™ or pJET vector respectively. Sequencing was conducted using the Hitachi 3730xl DNA Analyzer (Applied Biosystems) using the Big Dye Terminator v3.1 system that is based on the Sanger method.

2.12.2 Sequence Analysis and Construction of phylogenetic trees

Chromas[®] was used for editing sequences. Edited sequences were aligned using Bio-Edit with *MEGA 4* (Tamura et al., 2007) being used to construct phylogenetic trees. The trees were constructed based on the Maximum Composite Likelihood method and substitution model using Neighbour-Joining. The test of phylogeny was used based on 1000 bootstraps of replication and a pair-wise deletion of gaps. Substitutions included transitions and transversions, with the pattern among lineages being assumed to be homogeneous (Felsenstein, 1985; Saitou N & Nei M, 1987; Tamura *et al.*, 2007). The amplified DNA sequences were identified through homology searches using BLASTn against the NCBI non-redundant database.

Chapter 3: Diazotrophic diversity and expression in an Antarctic hypolithon

3.1 Introduction

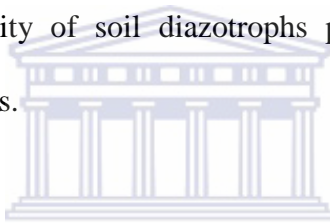
Soil microorganisms probably represent the world's greatest source of biological diversity, and play an important role in nutrient biogeochemical cycling (Yergeau *et al.*, 2007a). In the Antarctic, environmental conditions impose extreme pressures for the proliferation of diverse microbial communities (Wynn-Williams, 1989). It is not well known how these conditions affect microbial communities and, in particular, nutrient cycling which is primarily driven by such communities in Antarctica.

Nitrogen cycling in terrestrial ecosystems is an important biogeochemical cycle but remains poorly understood in most extreme environments. However, available evidence suggests that nitrogen is the main growth limiting factor in both high altitude and high latitude ecosystems (Mataloni *et al.*, 2000; Shaw & John, 2001). This implies that small changes in the nitrogen cycle, as predicted in climate change models, may result in disproportionately large shifts in the dynamics of the ecosystems (Barnard *et al.*, 2006). However, no study focusing on specific microbial communities implicated in nitrogen fixation exists.

Nitrogen fixation is the rate limiting step in the nitrogen cycle. Terrestrial diazotrophs are the main input source of nitrogen in primary-producing ecosystems (Poly *et al.*, 2001). In oligotrophic Antarctic biotopes, hypoliths are expected to be a significant input source of nitrogen.

All organisms capable of nitrogen fixation carry the *nifH* gene and it is thus considered to be an appropriate proxy for investigating organisms potentially capable of nitrogen fixation (Wang *et al.*, 1988; Zehr *et al.*, 2001; Zhang *et al.*, 2008).

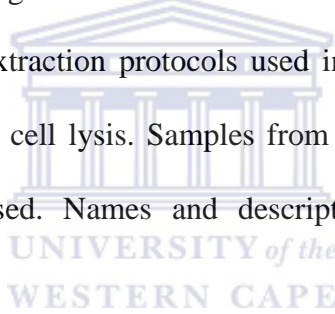
Recent studies focused on diazotrophic diversity in terrestrial systems have employed culture-independent approaches. These approaches present a more inclusive depiction of the diazotrophic community (Poly *et al.*, 2001). Techniques such as PCR-cloning (Zehr *et al.*, 2001), PCR-denaturing gradient gel electrophoresis (DGGE) (Coelho *et al.*, 2009) or PCR-restriction fragment length polymorphism (RFLP) (Poly *et al.*, 2000) have all been employed to show the diversity of soil diazotrophs potentially capable of nitrogen fixation in various environments.



In this study, a culture-independent approach was used to attempt to elucidate the diversity and gene expression of microorganisms potentially capable of nitrogen fixation in a hypolithic community. This is the first such study conducted in hyperarid Antarctic hypolithic communities. The structure of hypolithic diazotrophs was investigated through PCR-RFLP and sequencing was employed to establish the phylogenetic identity of these organisms.

3.2 DNA and RNA Extraction

In order to investigate the diversity and expression of functional genes in soil it is important to extract both DNA and RNA. Extraction protocols require optimization when applied to different soil types (Burgmann *et al.*, 2003; Miller *et al.*, 1999). This becomes increasingly important when attempting to obtain high quality DNA and intact RNA. Recovery of high yield DNA and RNA is crucial in order to ensure adequate accuracy in the representation of diversity (Miller *et al.*, 1999). Protocols which incorporate bead-beating (physical lysis) are assumed to introduce less bias as compared to chemical lysis methods since high concentrations of nucleic acids are recovered (Frey *et al.*, 2006). The nucleic acid extraction protocols used in this chapter all include bead-beating coupled with chemical cell lysis. Samples from hypolithic soil, non-translucent rock, and open soil were used. Names and descriptions of samples are attached (Appendix 2).



3.2.1 Simultaneous Extraction using CTAB lysis buffer

Soil previously stored in *RNAlater*® solution was used. High molecular weight DNA and RNA was obtained using the protocol described in Chapter 2. However, the amount of RNA recovered was very low (Figure 10) using 1g of soil. The initial amount of soil used for extraction was increased in an attempt to improve the final yield of RNA. However, these trials produced degraded DNA which was deemed to be unsuitable for downstream metagenomic analysis (Figure 11, using 3g of soil).

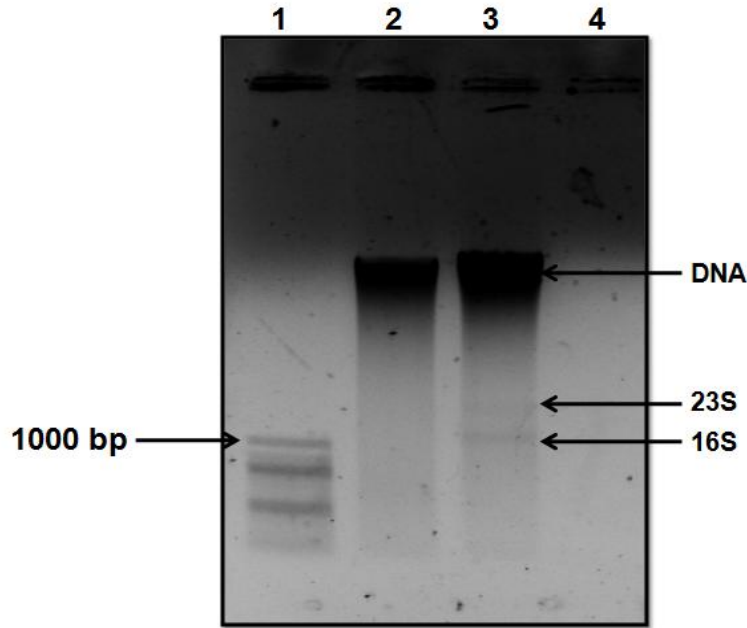


Figure 10: Simultaneous extractions of DNA and RNA from hypolith samples, (Miers Valley Hypolith ; (MVH)). Lane 1: RNA Marker (Fermentas) Lane 2: MVH 3A; Lane 3: MVH4A. 23S and 16S rRNA gene bands are only faintly visible. (Electrophoresis on 1% agarose gel); Lane 4: Blank

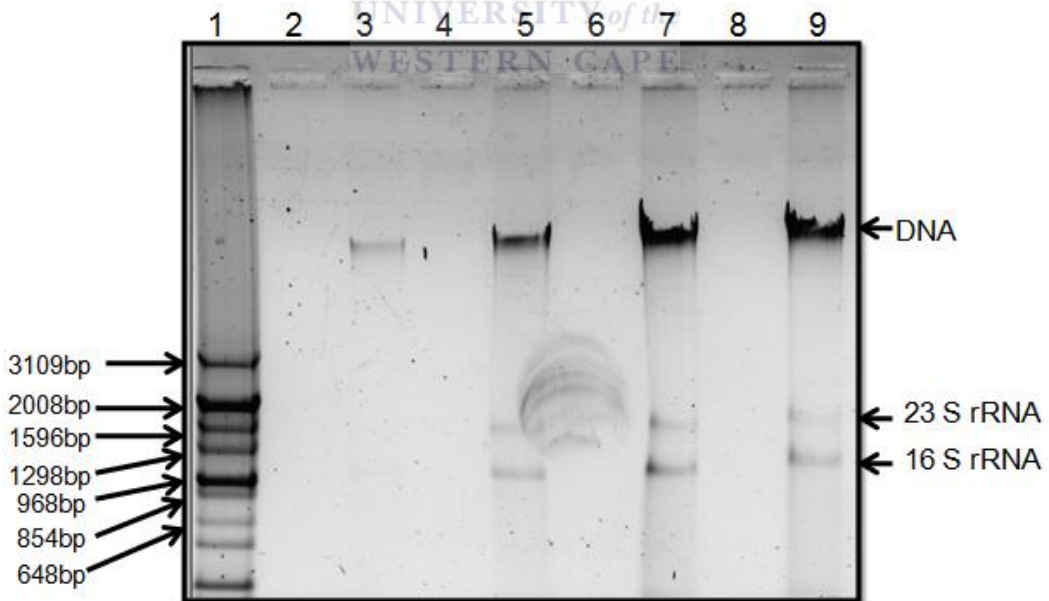


Figure 11: Simultaneous DNA/RNA extraction samples. Lane 1: Marker (IMBM marker), Lane 3: MVH 3A, Lane 5: MVH 4A, Lane 7: MVH 11A, Lane 9: MVH 12A (1% agarose gel electrophoresis).

Lanes 2, 4, 6, and 8: Blank.

3.2.3 MoBio Soil RNA kit

While the MoBio RNA kit resulted in good yields of total RNA, DNA was also co-extracted using this kit (Figure 12). The quality of RNA extracted was suitable for downstream applications since both the 23S rRNA gene and 16S rRNA gene bands were intact. In an extraction giving good quality yield, there is an approximate 2:1 ratio between the 23S rRNA band and the 16S rRNA band (as indicated by UV excitation on electrophoresis, based on visual inspection). There is evidence of degraded total RNA on the gel, although this is to be expected since soil contains both living- and dead organisms (Personal communication Katharine Coyne, University of Delaware).

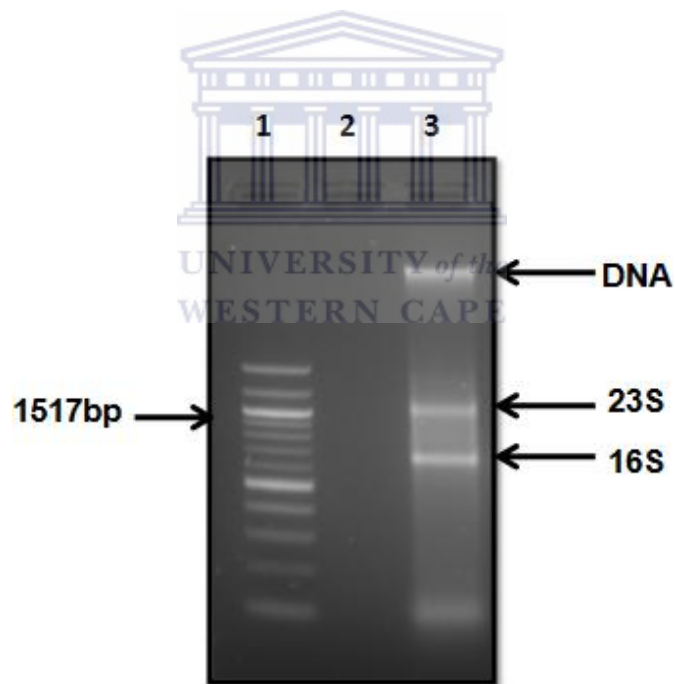


Figure 12: Electrophoresis of RNA extracted using the MoBio kit. Lane 1; 100 bp Marker (Fermentas). Lane 3: MVH 11A. Both 23S and 16S bands are clearly visible. Degraded RNA is visible at the bottom of the gel.

3.3 Expression of *nifH* genes

3.3.1 DNase treatment of simultaneous extractions

In order to obtain a profile of microorganisms actively expressing the *nifH* gene, mRNA free from contaminating DNA is required. Digestion with DNase I was carried out as described in Chapter 2 (Figure 13). However, the control 16S rRNA gene PCR (using the E9F and U1510R primer combination) designed to test for the presence of contaminating DNA was positive. Three successive DNase I treatments were carried out, but amplification was consistently observed in the control reactions which were carried out at different concentrations of the DNase I treated sample (Figure 14). No amplification was observed in the negative control (no template).

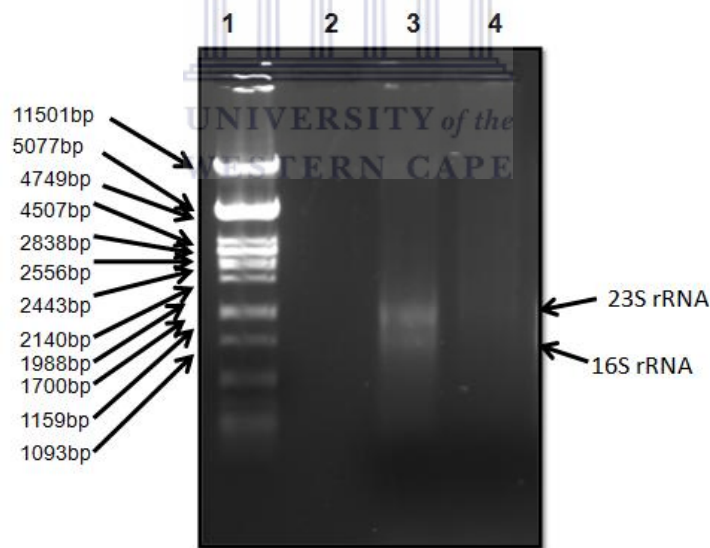


Figure 13: 4 DNase treatment of Simultaneous extracts (DNA and RNA).

Arrows indicate position of 23S rRNA gene and 16S rRNA gene bands. Lane 1: λ Pst Marker; Lane 2: left blank; Lane 3: DNA/RNA after 1 DNase I treatment, Lane 4: Lane 4 DNA/RNA after 3 treatments with DNase I (Electrophoresis on 1.5% agarose gel).

During each successive DNase treatment, large quantities of RNA as determined by quantification using Q-bit fluorometry analysis (data not shown) were lost and only small amounts of RNA were visible on the gel electrophoresis (Figure 13).

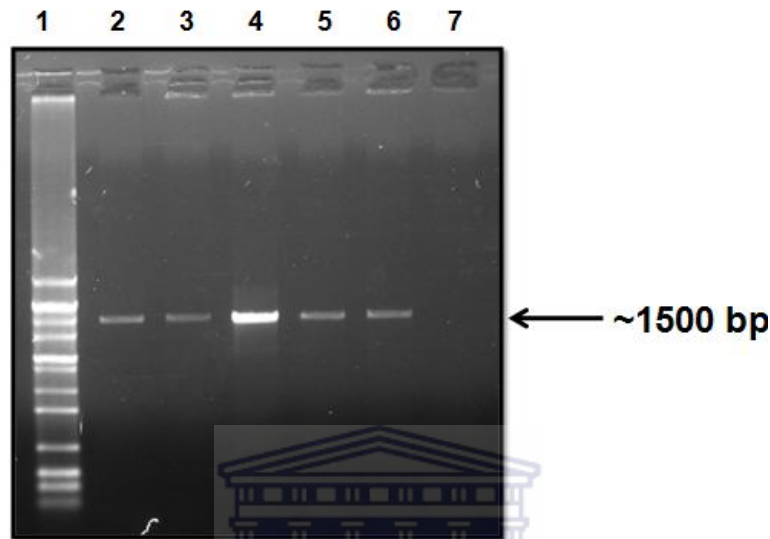
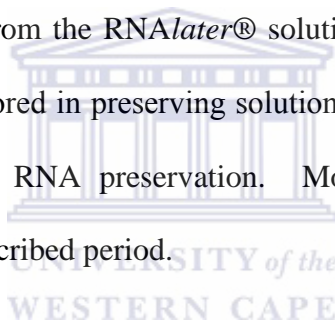


Figure 14: 16S rRNA gene PCR amplification from DNase treated total nucleic acid extracts indicating amplification at various concentrations. Lane 1: Marker (IMBM Marker), Lane 2 to Lane 6: different concentrations of template RNA (ranging from 10 ng/ul to 50 ng/ul). Amplification was most efficient with 30 ng/ul of template. Lane 7: Negative control (No template) (1% agarose gel electrophoresis).

RNA was not visible on the gel after the third DNase treatment suggesting that the residual level of RNA was below the UV- ethidium bromide detection limit (approximately less than 1ng/ul). The quantity of total RNA remaining was too low for successful RT-PCR amplification, and a decision was taken not to continue with attempts to obtain *nifH* gene expression profiles.

In order to investigate nitrogen cycling gene expression high quality total RNA is required. Previous studies were also unsuccessful in obtaining DNA-free RNA from oligotrophic hyperarid soils of the Dry Valleys (Personal communication Don Cowan, University of the Western Cape). The low environmental temperatures and low biomass may be barriers leading to low RNA levels. The methods used in this study require optimization for these soils. Moreover, the ability to obtain non-degraded RNA was hampered by storage of soil samples for prolonged periods at -80°C . Any attempt to recover usable RNA would require “freshly sampled soils”. Attempts to recover RNA may have been compromised by the large salt pellets which were co-extracted with the RNA. The salts are possibly from the RNeasy[®] solution. In future, samples intended for RNA analysis should be stored in preserving solution, such as LifeGuard[™] (MoBio) designed specifically for soil RNA preservation. Moreover, extractions should be attempted in the minimum prescribed period.



3.4 Diversity of *nifH* genes

3.4.1 RNase treatment of simultaneous extractions

Simultaneous DNA/RNA extracts were subjected to RNase treatment before downstream metagenomic analysis was carried out. DNA free from contaminating RNA was obtained and used for further analysis (Figure 15). The DNA extracted was of good quality and high molecular weight, approximately 11501bp. No evidence of RNA contamination was visible on the gel, nor was it detected by fluorometry (data not shown).

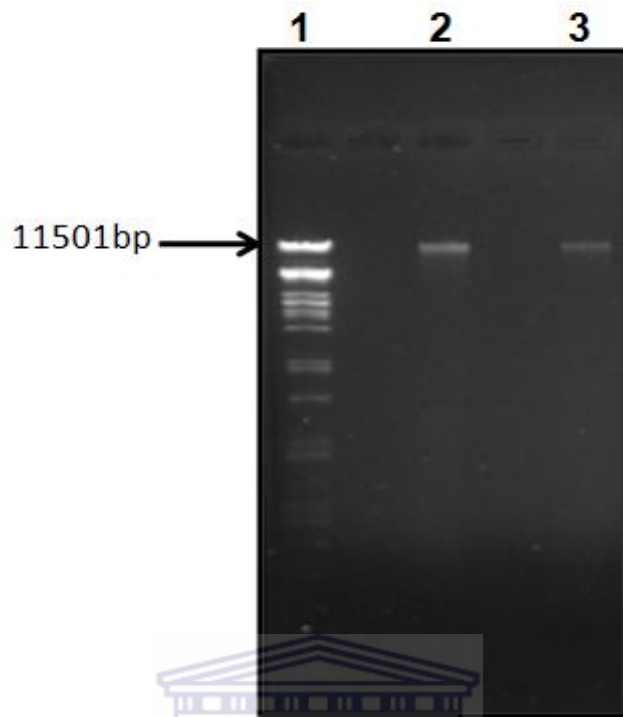


Figure 15: RNase treatment of DNA/RNA extracts to remove RNA

Lane 1: λ -Pst Marker, Lane 2: MVH 11A, Lane 3: MVH 12A (1% agarose gel).

3.4.2 PCR amplification of the 16S rRNA gene with DGGE primers

PCR was carried out as described in Chapter 2, section 2.6. 16S rRNA gene PCR, using the universal bacterial DGGE primers 341FGC and 534R was successful for hypolithic samples, yielding an approximate 200 bp PCR product (Figure 16). Amplification products were then analysed using DGGE (section 3.4.3).

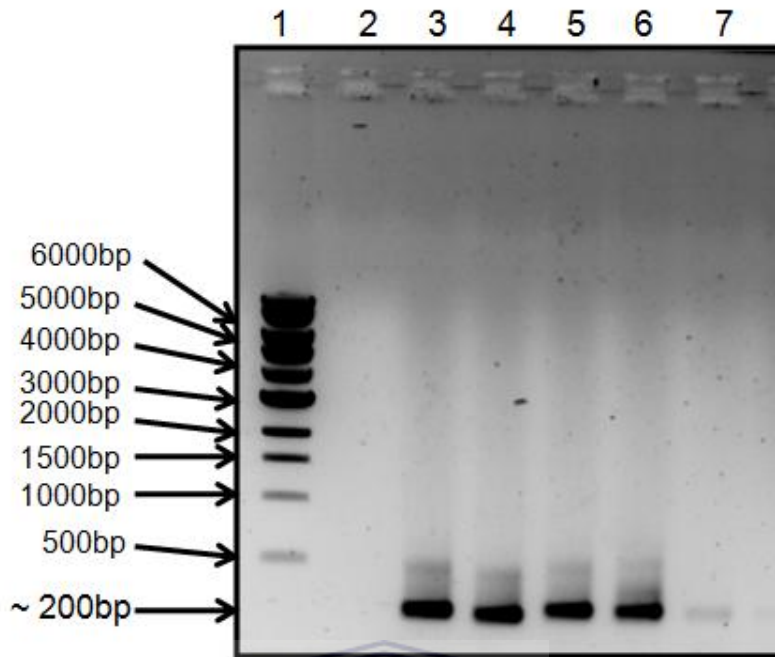
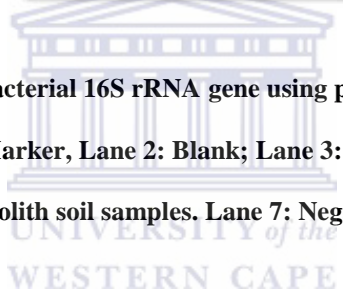


Figure 16: PCR amplification of bacterial 16S rRNA gene using primers 341 F-GC and 534R on 1% agarose gel. Lane 1: Biolabs 1kb Marker, Lane 2: Blank; Lane 3: *E. coli* Positive Control, Lane 4 -6: Various hypolith soil samples. Lane 7: Negative control



3.4.3 DGGE

DGGE analysis was used to obtain an overview of the bacterial diversity in hypoliths; open soil and non-translucent rock subsoil samples (Figure 17). It was seen that all samples contained significant bacterial diversity (with each band representing at least one microorganism). Many of the observed phlotypes were common to every sample. Open soil and non-translucent rock community soil were observed to lack certain phlotypes (shown in red arrows) which were dominant across all amplified hypolithic communities. Amplification of *E. coli* (from *Taq* polymerase) was observed in all samples. This is due to use of native *Taq* polymerase (IMBM Lab *Taq*). The corresponding sequenced band

(Marker E) was previously observed in previous studies conducted in this laboratory (Keyster, 2007; Khan, 2008).

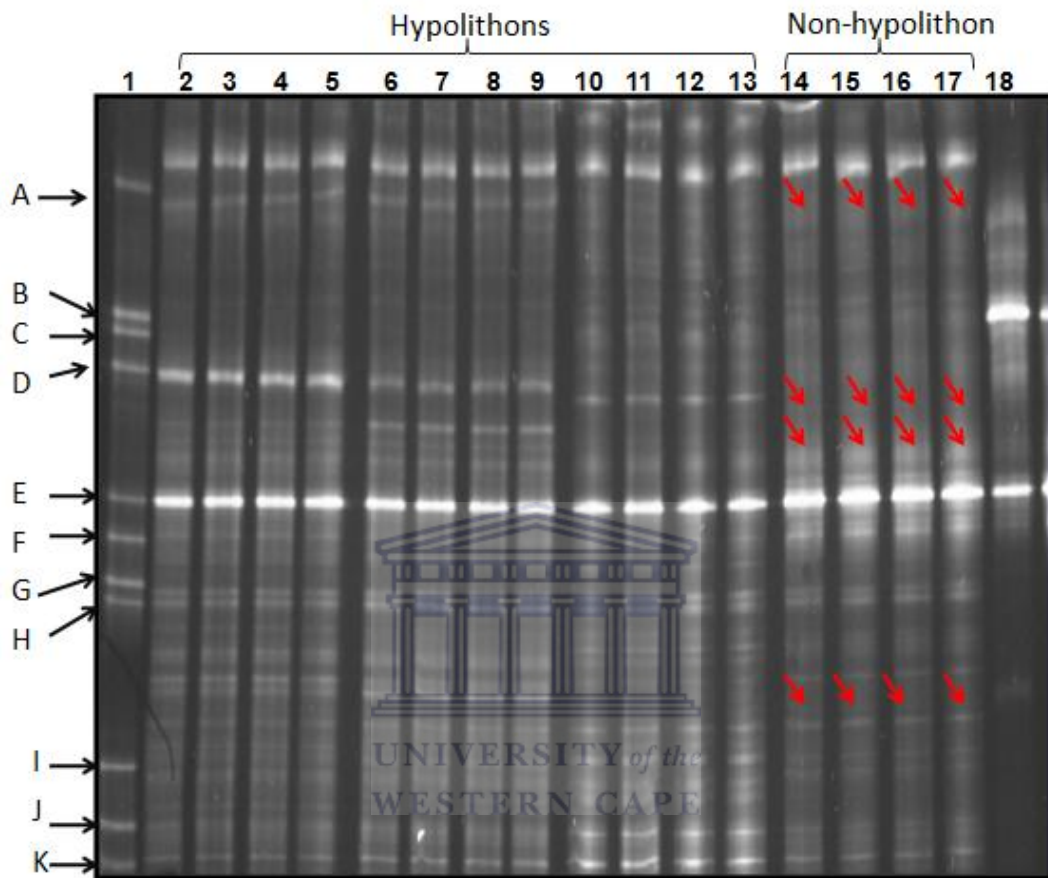


Figure 17: DGGE Analysis of Antarctic Dry Valley hypolith samples, open soil and non-translucent rock soil. Duplicate extractions were carried out for each sample before separation on DGGE. Lane 1 DGGE Marker, Lane 2 and 3: MVH 3A, Lane 4 and 5: MVH 4A, Lane 6 and 7: MVH 5A, Lane 8 and 9: MVH 9A, Lane 10 and 11: MVH 11A, Lane 12 and 13: MVH 12A, Lane 14 and 15, MVH 3C (Open soil), Lane 16 and 17: MVH 3B (non-translucent rock), Lane 18 and 19: *Nostoc sp.* DNA.

Analysis of DGGE profiles and banding patterns was performed using GelCompar® II, version 5.0 (Applied Maths) as described by the manufacturer. A similarity dendrogram was generated as described in Chapter 2, based on the Pearson correlation coefficient

together with the Unweighted Pair Group Method using Arithmetic averages (UPGMA). The dendrogram shows two main clusters. A hypolithic cluster formed by various hypolith samples, and a cluster consisting of open soil and non-translucent rock subsoil phylotypes (Figure 18). This separation into distinct clusters is supported by high bootstrap values (Figure 18) as indicated by the similarity dendrogram.

A Multi-dimensional scaling (MDS) model was also generated (Figure 19). This model further validates the similarity of hypolithic communities based on band matching. All hypolith samples formed clusters, while the open soil and non-translucent rock samples assembled together.



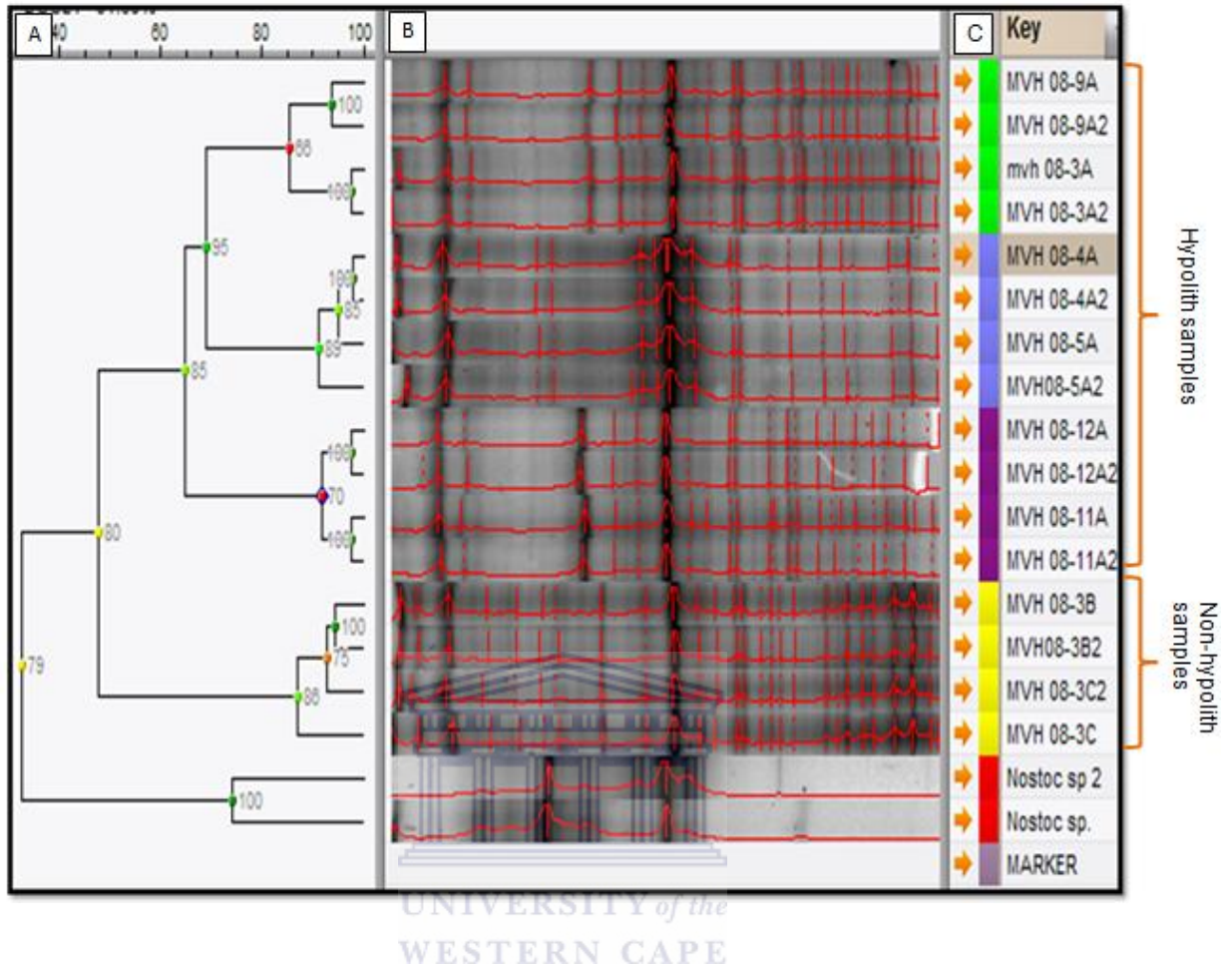


Figure 18: Gelcompar Screen-shot of (A) Similarity dendrogram generated from DGGE profile. (B) DGGE profile with electrophoresis peaks from bands. (C) Sample sorted according to hierarchical clustering from bands.

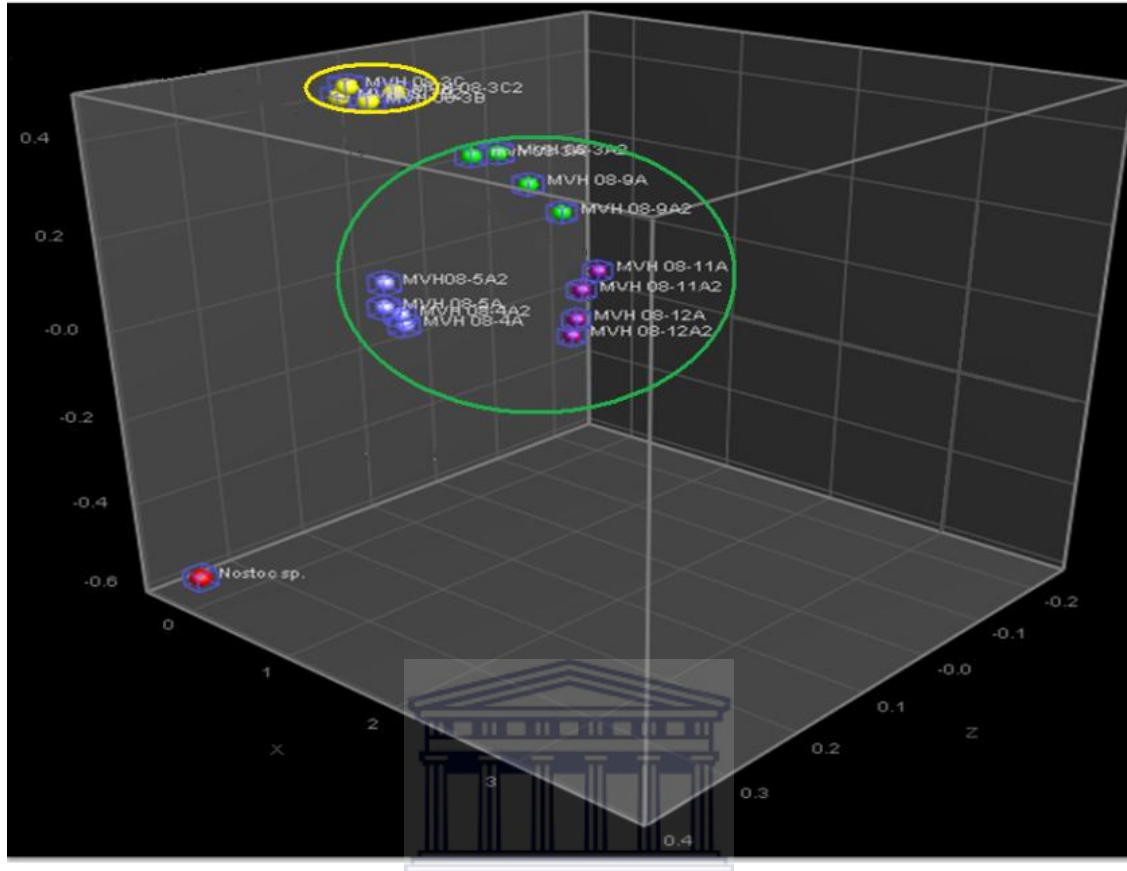


Figure 19: MDS ordinations showing relationship between terrestrial habitats based on the presence of bacterial genospecies. Hypolithic communities are closely related to each other, as indicated by observed clusters.

Both the dendrogram (Figure 18) and MDS analysis (Figure 19) confirm that hypolithic communities are similar to each other. The *Nostoc* sp. (positive control) formed a separate cluster. The established cluster positions confirm homogeneity in terms of banding patterns. Therefore, based on banding patterns a single hypolithon can be used as a representative.

3.4.4 PCR with *nifH* primers

Previously designed degenerate *nifH* primers (PolF and PolR) were used to amplify the conserved *nifH* gene region in diazotrophic species. A 360 bp *nifH* fragment was successfully amplified under the conditions described by Poly *et al.*, (2000) (Figure 20). Due to the similarity in bacterial community structure in different hypolithic communities (section 3.4.3), a single hypolithon was chosen for library construction and further downstream analysis.

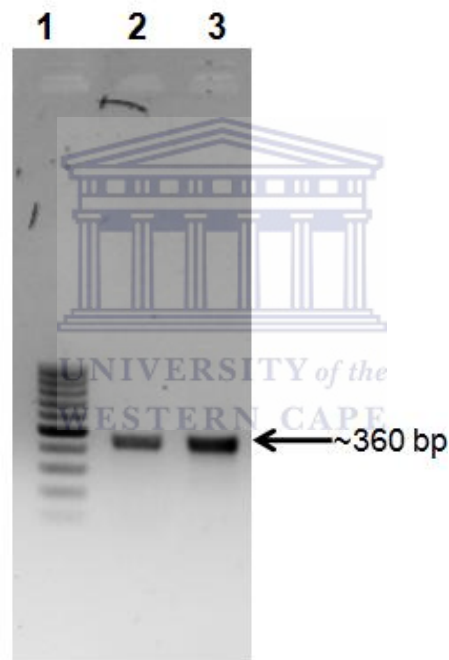


Figure 20: Amplification of the *nifH* gene fragment from hypolithic samples. Lane 1: *Azotobacter vinelandii*, Lane 2: MVH 11A Lane 3: 100bp Marker (Fermentas) (1% Agarose gel of *nifH* amplified fragment)

3.4.5 *nifH* Clone library

nifH amplicons from the hypolith sample MVH11A were cloned as described in Chapter 2. PCR using M13 vector primers was used to confirm the presence of correct sized insert (Figure 21). Over 80% of clones contained the correct size fragment. A total of 124 clones were analyzed by RFLP analysis as a means of de-replicating similar clones.

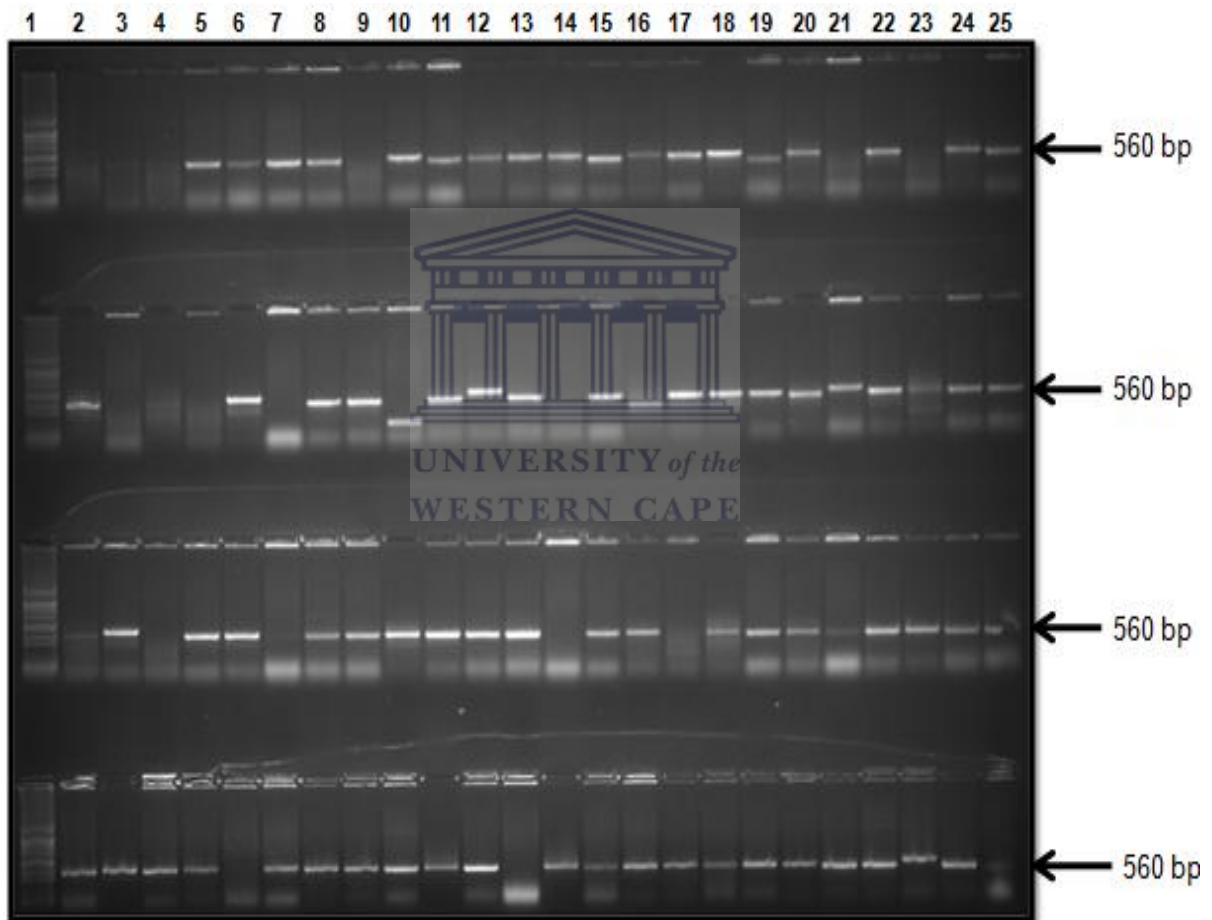


Figure 21: Colony PCR indicating the presence of *nifH* fragments in over 80% of screened clones (2.5% agarose gel).

3.4.6 RFLP analysis of clones

RFLP analysis was performed on M13 PCR products obtained from the clone library. An online program (WatCut) was used for selecting the restriction enzyme to be used for RFLP analysis (<http://watcut.uwaterloo.ca/>). *Hae*III was selected for all restriction analyses. Digested products were visually classified into operational taxonomic units (OTU), and formed the basis for further analysis (Figure 22). Eighteen unique patterns were observed in the library from the 124 clones analysed.



Figure 22: *Hae*III RFLP analysis of M13 PCR amplicons obtained from *nifH* library clones (2.5% agarose gel)

3.4.6.1 Richness Estimation of RFLP groups

Phylotype- richness estimators are used to judge the phylogenetic diversity of a library.

The total number of putative *nifH* RFLP groups from the library was estimated using the

S_{Chao1} and S_{ACE} programs available online

(<http://www.aslo.org/lomethods/free/2004/0114a.html>).

S_{Chao1} is defined as a non-parametric estimator based on the mark-recapture technique

(Chao, 1984), calculated using the following formulae;

$$S_{\text{Chao1}} = S_{\text{obs}} + \frac{F_1^2}{2(F_2 + 1)} - \frac{F_1 F_2}{2(F_2 + 1)^2}$$

Where;

S_{obs} is the number of phylotypes in the library

F_1 and F_2 are the number of phylotypes occurring as singletons or doubletons, respectively.

S_{ACE} is a coverage based estimator defined as follows;

$$S_{\text{ACE}} = S_{\text{abund}} + \frac{S_{\text{rare}}}{C_{\text{ACE}}} + \frac{F_1}{C_{\text{ACE}}} \gamma_{\text{ACE}}^2$$

Where F_1 is the number of phylotypes occurring more than once in a library,

S_{rare} is the number of phylotypes occurring 10 or fewer times, and S_{abund} is the number

occurring more than 10 times. γ_{ACE^2} is the coefficient of variation of the F_1 's. C_{ACE} is a

sample coverage estimate defined as the proportion of individuals in relatively rare

phylotypes (<10 clones) occurring more than once in a library (Chao *et al.*, 1993)

Table 5 summarizes the calculated S_{ACE} and S_{Chao1} values obtained after fitting data into the online model. The ratio of S_{ACE} and S_{Chao1} is close to 1 in both instances, indicating that the phlotypes observed were accurate (Table 5). The profile of each curve generated from this model can be used as an indicator of library coverage (Figure 23). In a well sampled library, there should be an agreement in the shape of the curves A and C with B and D (Figure 23) (Kemp & Aller, 2004). All the graphs indicate that the clone library was sufficiently covered. This is typical in libraries where most phlotypes may occur only once or twice (Kemp & Aller, 2004).

Table 5: Richness Estimation Prediction model for *nifH* Clone library

Number of clones in library	124
Number of phlotypes observed	18
Predicted value of S_{ACE}	18.73
Predicted value of S_{Chao1}	18.40
Observed phlotypes / predicted S_{ACE}	0.96
Observed phlotypes / predicted S_{Chao1}	0.98

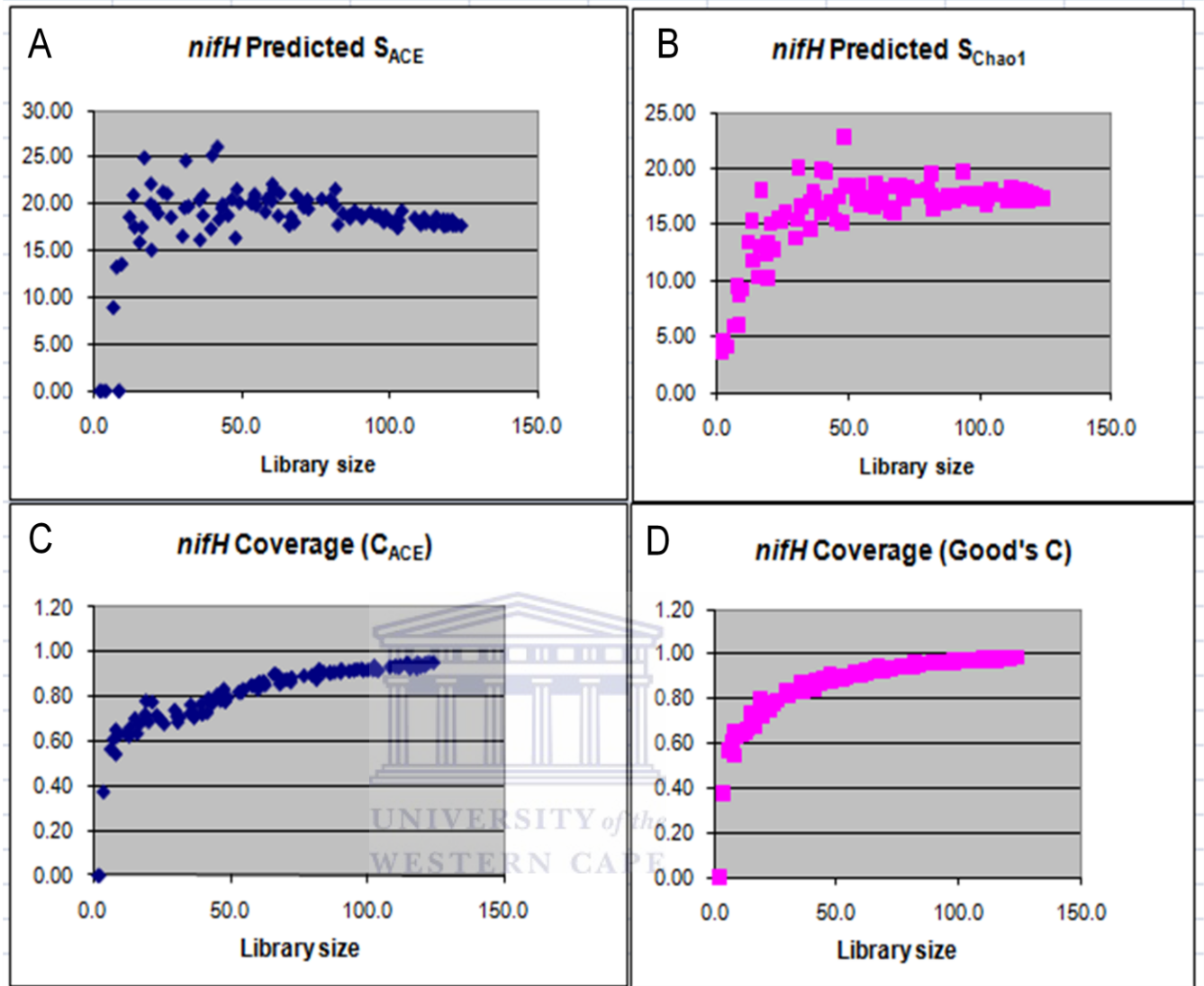


Figure 23: Library richness estimation. In a well sampled library there is concurrence in the shape of
pane A and C with B and D.

3.4.7 Phylogenetic Analysis

The phylogenetic affiliation of each phylotype was identified based on homology searches to similar sequences in the NCBI database (<http://blast.ncbi.nlm.nih.gov/Blast.cgi#>). Nucleotide sequence alignments indicate a high level of similarity of sequenced clones to known database *nifH* sequences (Appendix 2). A representative of each OTU group was sequenced. The closely related phlotypes from the database were identified and their respective nucleotide sequences were used as positive controls for alignments based on sequence identities. The alignment was employed for construction of phylogenetic trees based on the similarities of the nucleotide sequences. The tree was used to determine the phylogenetic positions of the phlotypes in relation to the known sequences. *nifH* PCR clones clustered with known cyanobacteria, and Proteobacteria. The majority of clones were positioned with *nifH* sequences of known cyanobacteria, indicating a dominance of these phototrophic organisms in hypolithons (Appendix 2). Overall, clones attained had a low percentage similarity (<90%) to *nifH* sequences of previously cultured strains deposited on the NCBI database. This low similarity indicates the potential presence of novel sequence features.

In order to infer phylogenetic affiliation of *nifH* sequences generated in this study both cultured- and uncultured database clone representatives were included for phylogenetic analysis. Relationships based on nucleotide sequences attained are discussed below. Four distinct clusters were identified from the tree topology (Figure 24). Most *nifH* clones were highly similar (>95%) to other uncultured clones from Antarctica marine environments (Jungblut & Neilan, 2009).

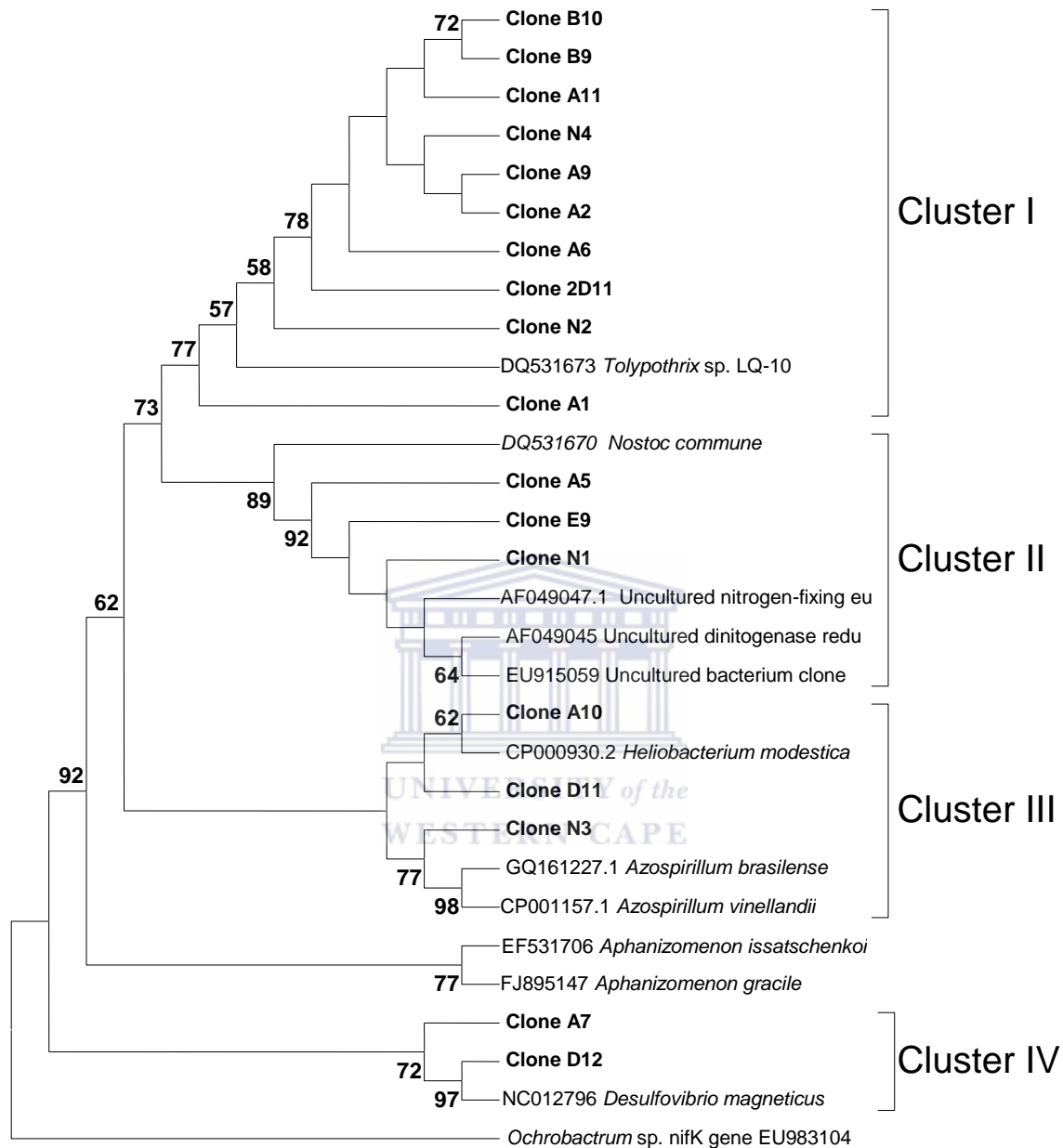
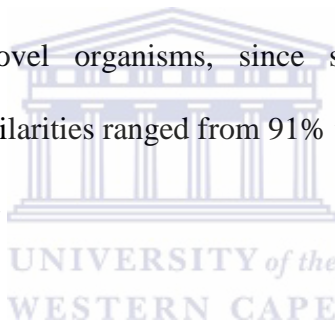


Figure 24: The evolutionary history of *nifH* sequences was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 9.11 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base.

Cluster I

Clones **B10, B9, A11, N4, A9, A2, A6, 2D11, N2** and **A1** formed this cluster. The clones were found to be highly similar to the *nifH* sequences of Nostocales cyanobacteria. The highest BLAST n match to a cultured representative identified from the tree topology was *Tolypothrix* sp. (DQ 531673), a member of the Nostocales family. All clones had higher BLAST n similarity to previously uncultured nitrogen fixing bacterium (Appendix 2).

The clones obtained in this study were related to uncultured *nifH* clones previously identified from Antarctica (Appendix 2). The BLAST n results indicate the *nifH* clones may represent potentially novel organisms, since similarity to known database cyanobacteria is fairly low, similarities ranged from 91% to 93%. This cluster represented the majority of obtained clones.

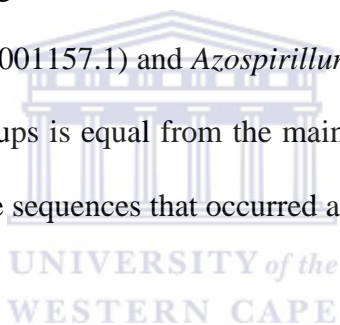


Cluster II

Cluster II was formed by Clones **A5, N1** and **E9**. This second cluster of hypolith *nifH* clones has sequences similar to previously cultured cyanobacteria. BLAST n analysis reveals a close affinity to *Nostoc commune* (DQ 531670). According to BLAST n results, all three clones have 95% sequence similarity to *Nostoc* sp.. Clones affiliated with this cluster had >98% similarity to uncultured *nifH* clones of Antarctic origin found in a published study focused on marine derived *nifH* clones (Jungblut & Neilan, 2009) (Appendix 2).

Cluster III

Three RFLP types are represented in this cluster (Clones **A10**, **D11** and **N3**). The clones were found to have a high sequence similarity to members of the Firmicutes and Proteobacteria, using BLASTn alignment. Representative Clone **A10** *nifH* sequence was 87% similar to the Firmicute *Heliobacterium modesticaldum* (CP000930.2); however in the tree topology the cluster between the two occurred only 62% of the time. BLASTn results indicate that Clone **D11** had a 99% similarity to *Azospirillum vinellandii* (CP001157.1). According to the phylogenetic tree, Clone **D11** has greater sequence similarity to Clone **A10** forming a sister clade, which diverges in lineage 77% of the time to *Azospirillum vinellandii* (CP001157.1) and *Azospirillum brasilense* (CP GQ161226.1). The distance of both sister groups is equal from the main node which possibly indicates mutations in specific nucleotide sequences that occurred at the same evolutionary time.



Cluster IV

Clones **A7** and **D12** both form Cluster IV. These two clones cluster together 72% of the time, based on bootstrap values inferred from the tree. The closest represented cultured *nifH* sequence is related to *Desulfovibrio magueiticus* (NC012796). In terms of aligned nucleotide sequences Clone D12 is most related, as seen from the tree topology.

3.5 Discussion

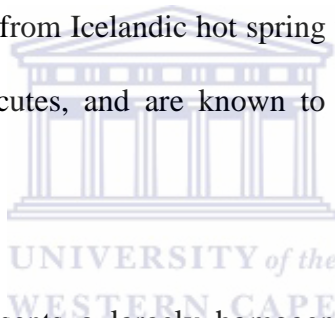
PCR is widely acknowledged to have inherent biases when used to assess microbial diversity (Farrelly *et al.*, 1995). PCR primers have been reported to overlook half the microbial diversity in a sample. This may be due to the relative inefficiency of extraction techniques (Hong *et al.*, 2009). However, these techniques are more effective in elucidating the extent of microbial diversity than culturing methods.

Previous studies, focused on Antarctic soil, have found that microbial diversity in these soils has previously been underestimated (Cowan *et al.*, 2002; Smith *et al.*, 2006). A range of biotopes have been found to harbour a large diversity of microorganisms. Among these biotopes are hypolithic communities which have been found to be dominated by microorganisms known to be important contributors to nutrient cycles (Khan, 2008; Pointing *et al.*, 2009). Previous characterization of hypolithic communities were based on the 16S rRNA gene. It was important therefore to undertake a baseline study focusing on the diversity of key anabolic genes. In this chapter the *nifH* gene was used to investigate the diversity of putative nitrogen fixing microorganisms in hypolithons.

DGGE analysis confirmed a high level of bacterial diversity in Antarctic terrestrial ecosystems. Hypolithic communities were observed to have a dominance of certain phylotypes, not found in open soil or non-translucent rock samples. MDS and cluster analysis confirmed this observation. DGGE analysis suggests that habitat specific factors could possibly shape the microbial diversity in hypolithons.

Amplified *nifH* gene markers belonging to Cluster I and Cluster II (Figure 24), were closely affiliated to known cyanobacteria. This was not surprising as cyanobacteria have been identified in hypolith and other rock associated communities through the use of 16S rRNA gene markers (Pointing *et al.*, 2009). The majority of these were affiliated to uncultured environmental samples with a close similarity to *Nostocales*.

Clusters III, IV, V and V consist of sequences with affinities to an assortment of proteobacteria and Firmicutes genera. Interestingly, Cluster III consists of two clones closely related to *Heliobacterium modesticaldum* strain Ice1, a unique anoxygenic phototroph which was isolated from Icelandic hot spring volcanic soils. Heliobacteria are the only photosynthetic Firmicutes, and are known to carry out fix nitrogen fixation (Sattley *et al.*, 2008).



Overall, the clone library presents a largely homogenous composition of *nifH* gene sequences. This reduced heterogeneity may be explained by environmental stressors, which may lead to a gradual homogenization of the community structure (Doran *et al.*, 2002).

Chapter 4: Phylogenetic diversity of autotrophic carbon sequestering genes in Antarctic hypolith communities

4.1 Introduction

Soil is a complex and dynamic biotope (Nannipieri *et al.*, 2003). The composition of microbial communities in this biological system remains poorly understood. Conservative estimates are that less than 1% of microorganisms have been cultured (Torsvik & Øvreås, 2002; Yergeau *et al.*, 2007b). There are two main justifications for intensifying studies focused on soil biodiversity. The first is that characterization of diversity will facilitate an understanding of the ecosystem and thus enable manipulation of its workings. The second is that the ability of ecosystems to withstand serious disturbances may depend, in part, on the diversity of the biological system (Torsvik & Øvreås, 2002), and we may therefore be able to predict the effects of environmental impacts such as global warming.

Antarctic terrestrial ecosystems contain low biomass. An absence of higher plants implies that the bulk functioning of soil systems is reliant on microorganisms. It is therefore important to determine the composition of microbial communities potentially implicated in important nutrient cycles. Antarctic hypoliths, important hyperarid desert communities, represent microbial populations which may be crucial for carbon input into the system.

Carbon dioxide sequestration in soil communities is chiefly mediated by microorganisms (Lal, 2008). Previous characterization of these communities based on 16S rRNA gene sequencing suggests the presence of autotrophic groups (Khan, 2008; Pointing *et al.*, 2009). However, no investigation has determined the functional composition based on the genes involved in carbon fixation.

The most important enzyme in CO₂ fixation is ribulose-1, 5-bisphosphate carboxylase oxygenase (RubisCO). The large subunit of RubisCO is encoded by the *cbbL* gene and is widely considered a functional marker for carbon fixation (Elsaied & Naganuma, 2001; Sinigalliano *et al.*, 2003; Tourova *et al.*, 2006). Phylogenetic sequence analysis suggests that the *cbbL* gene may be divided into two major groups, the green-like and the red-like groups (Chapter 1, section 1.4.2).

This study attempts to determine, for the first time, the presence of these genes in highly specialized Antarctic hypolithons. Using a previously designed primer set, the *cbbL* gene was used to probe for the presence of green-like and red-like bacterial RubisCO genes. As a baseline study, this investigation could help establish how perturbations in environmental parameters affect the composition of these communities.

4.2 Results and Discussion

4.2.1 DNA extraction from environmental samples

For this study DNA was isolated from hypolithic communities using a Modified protocol first described by Miller and colleagues (Figure 25) (1999) (Chapter 2). High quality, non-degraded metagenomic DNA was extracted. The extracted DNA was purified as described in section 2.4.4 in Chapter 2.

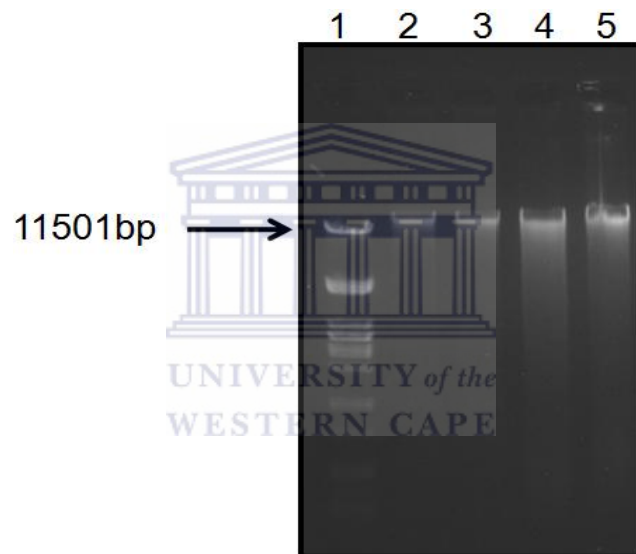


Figure 25: Agarose gel (1%) electrophoresis of DNA extractions. Lane 1: λ - Pst Marker, Lane 2 and Lane 3: MVH 3A, Lane 4 and 5: Control soil (UWC)

4.2.2 Amplification of *cbbL* genes

Amplification of the RubisCO genes from hypolith DNA via PCR was performed with the respective primer pairs (Table 3, Chapter 2) (Figure 26). Amplification with the specific primer combinations for red-like and green-like *cbbL* genes yielded the expected products of 800 bp and 1,100 bp respectively.

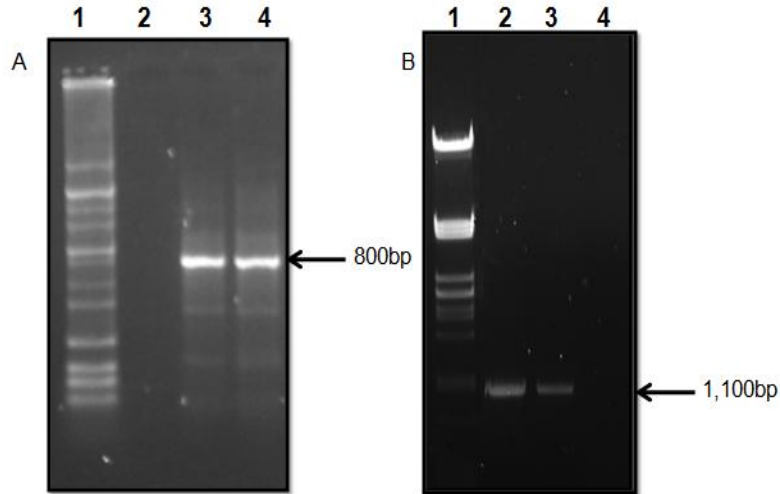


Figure 26: Agarose gel (1%) electrophoresis of amplified *cbbL* amplicons. (A) Lane 1: Marker (IMBM marker), Lane 2: Negative control, Lane 3: *Bacillus licheniformis* (Positive control), (B) Lane 1: λ Pst Marker:, Lane 2: Positive Control Lane 2: MVH 3A sample Lane 3: Negative control.

4.2.3 *cbbL* Clone library

4.2.3.1 *cbbL* Red-like

The 800 bp amplicon obtained using red-like *cbbL* primers was used to construct a clone library. A total of 185 of 236 clones from the library showed the correct fragment size, and were used for downstream analysis.



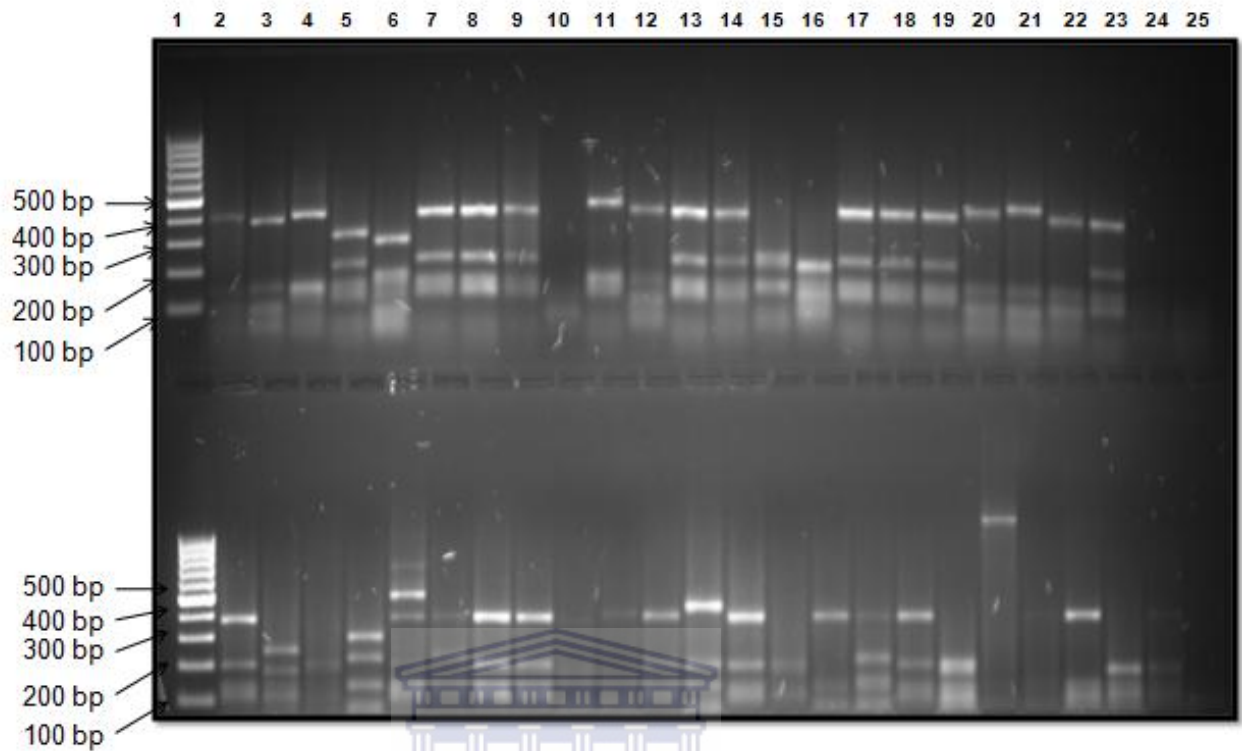


Figure 27: 2.5% Agarose gel of *cbbL* red-like RFLP digests with *Hae*III. Lane 1: 100bp DNA ladder (Fermentas), Lane 2-25: Selected M13 amplicons.

4.2.3.2 *cbbL* Green-like

PCR was performed with Phusion Taq® polymerase yielding blunt-end amplicons. Correct sized amplicons (1,100 bp) were cloned into pJET vector (to facilitate blunt end cloning). Fifty clones were screened for the presence of the insert. After RFLP analysis a total of 6 different restriction patterns were identified. The *cbbL* green-like clones represented here are not deemed to represent the complete diversity of organisms harbouring *cbbL* green-like genes, since exhaustive screening was not undertaken.

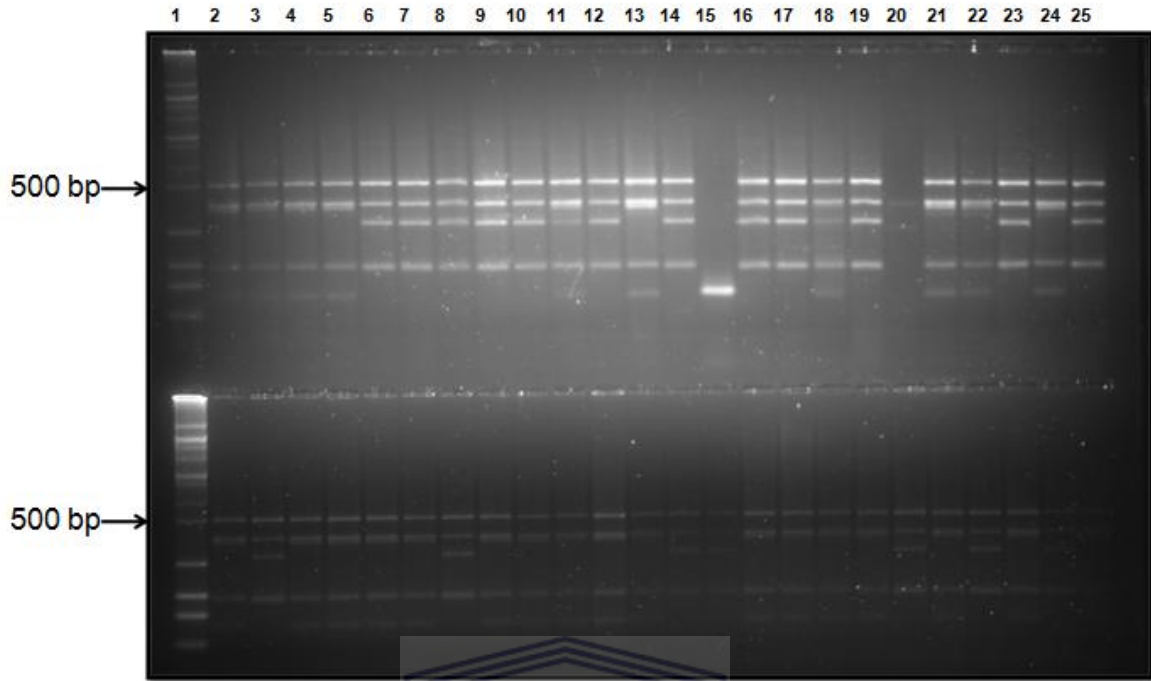


Figure 28: 2.5% Agarose gel of random amplicons digested with *AluI*. Lane 1: IMBM lab marker; Lanes 2-25: Selected restriction digests products.

UNIVERSITY of the
WESTERN CAPE

4.2.4 Diversity indices of clone libraries

4.2.4.1 *cbbL* Red-like

The restriction profiles were compared and manually grouped into 20 phylotypes. In order to assess if the clone library had sufficient coverage, the web interface constructed by Kemp and Aller (2004) was used. Phylotype richness (S_{Chao1} estimator) was compared against sampling effort (library size). The obtained curves showed that richness estimates reached a stable asymptotic value indicating that the number of clones analyzed was sufficient to yield a reliable estimate of phylotype richness. Richness estimates were slightly greater than the actual number of phylotypes (Table 6). Calculations of C_{ACE} and

Good's C (Figure 29 panel D) confirmed that clone library was large enough to capture the diversity of *cbbL* red-like genes. Coverage was estimated to be 0.97 and 0.96 using the S_{ACE} and S_{Chao1} estimators, respectively. Representatives from each of the OTUs were sequenced. Due to the limited number of OTUs no extensive sampling was undertaken.

Table 6: Richness Estimation Prediction model for *cbbL* red-like clone library.

Number of clones in library	185
Number of phylotypes observed	20
Predicted value of S_{ACE}	20.59
Predicted value of S_{Chao1}	20.64
Observed phylotypes / predicted S_{ACE}	0.97
Observed phylotypes / predicted S_{Chao1}	0.96

UNIVERSITY of the
WESTERN CAPE

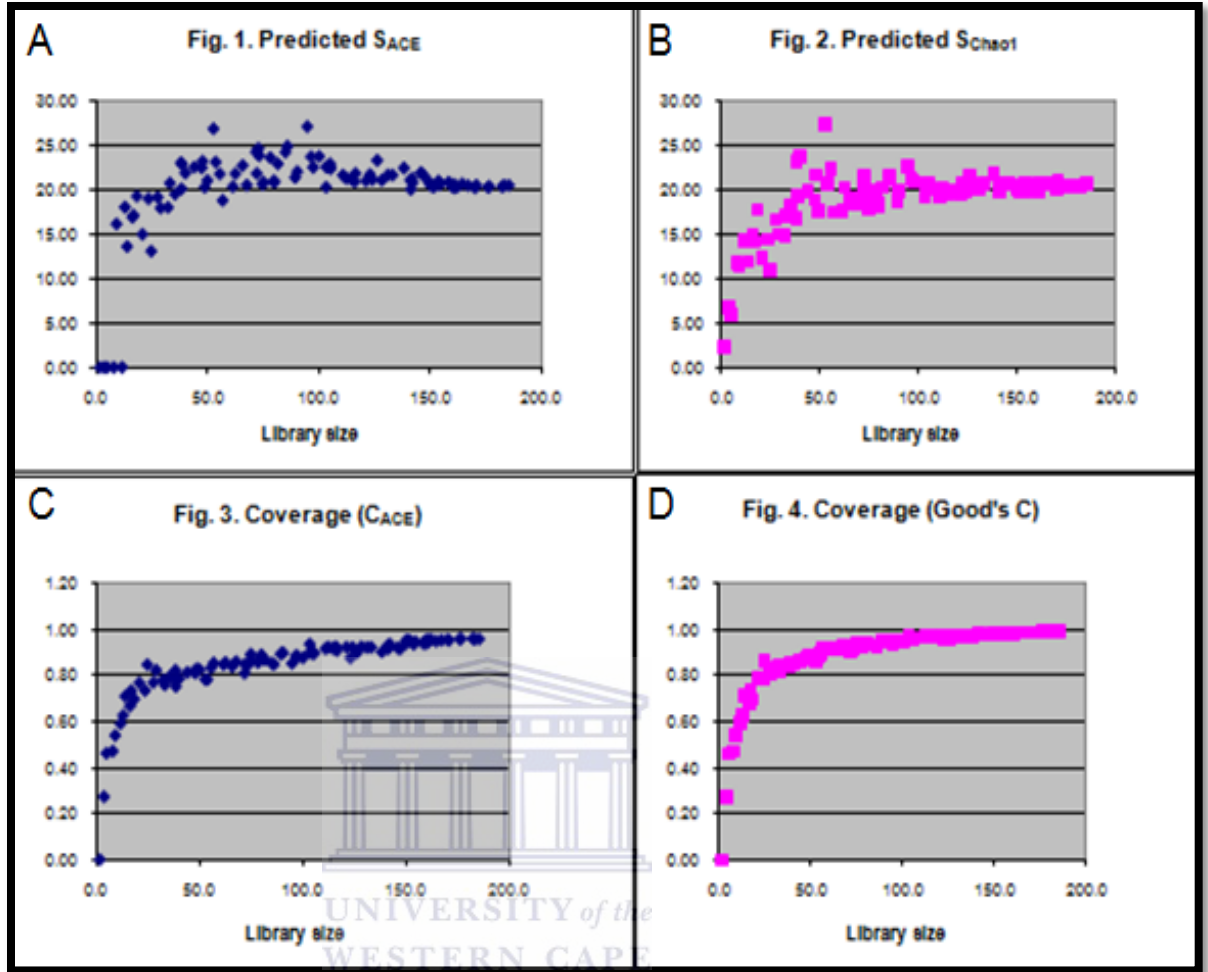


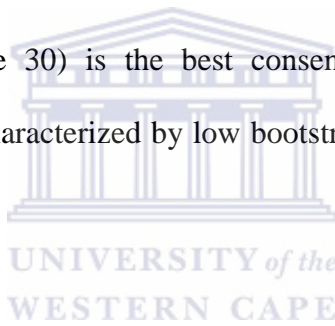
Figure 29: Library richness estimation for *cbbL* red-like libraries. In a well sampled library there is concurrence in the shape of panel A and C with B and D (Kemp & Aller, 2004).

4.3 Phylogenetic analysis based on *cbbL* sequence data

4.3.1 *cbbL* red-like

Representative clones from the *cbbL* red-like library were sequenced. Sequence similarities ranged from 85% – 96% (Appendix 3). The sequences formed three distinct clusters (Figure 30). Phylogenetic analysis revealed that all sequences were related to known *cbbL* gene sequences. A total of 20 representative clones were sequenced, one was found to be a product of primer mismatch and was discarded from subsequent analysis. Phylogenetic inference was based from nucleic acid alignments.

The phylogenetic tree (Figure 30) is the best consensus tree which was obtained; although certain nodes were characterized by low bootstraps from the tree topology four main clusters were identified;



Cluster I:

Clones **RH4, RH3, RG8, RH12, R2A1, R3B3, RD5, RF3, RF11, RH9, RH3, RF10, RD6, RD9, RC4, and RD3** were representatives forming this cluster. The clones had *cbbL* red-like sequences highly similar to that of *Mycobacterium sp.* (EU026272.1). Mycobacteria, and actinobacteria, are surprisingly closely related to this clade which contains clones with BLASTn similarities close to cultured *cbbL* red-like sequences of α -, and β -Proteobacteria.

Cluster II:

Clone 4A4 formed the second cluster. Based on the tree topology **Clone 4A4** forms a solitary cluster. Bootstrap support for formation of this cluster is 60%. From the tree topology a high similarity is implied with Cluster I clones. This suggests a close relation between the *cbbL* red-like sequences of the two clusters.

Cluster III:

Representative clones **3A3R** and **3A6R** were members of this cluster. From the tree topology the most common sister clade, in evolutionary terms, is a clade of uncultured *cbbL* red-like genes.



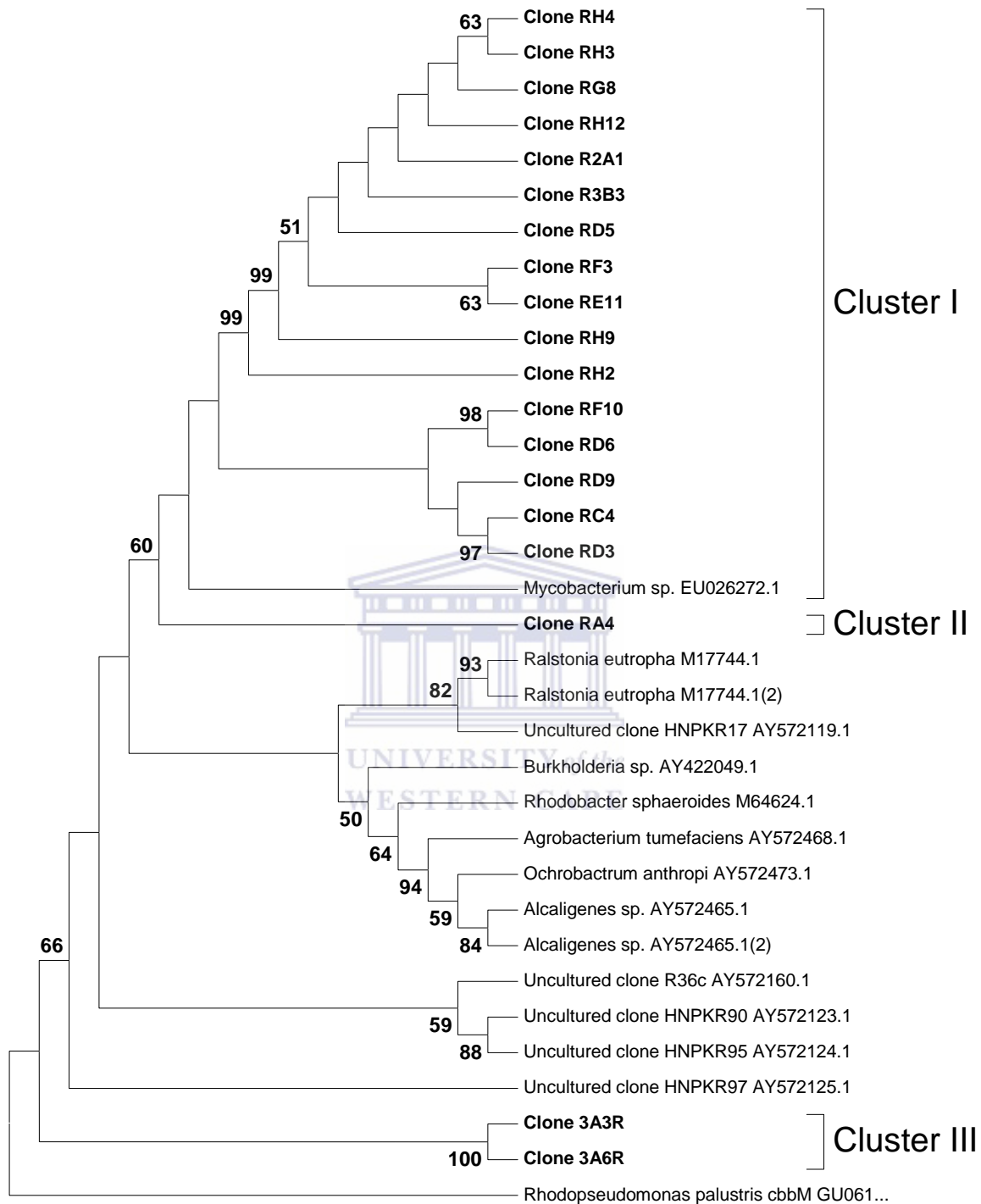


Figure 30: The evolutionary history was inferred for *cbbL* red-like sequences using the Neighbor-Joining method. evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total of 371 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4.

4.3.2 *cbbL* green-like

The restriction profiles were compared and manually grouped into 6 phlotypes. The highest BLAST n results were used as controls for construction of the phylogenetic tree.

According to the tree topology, three main clusters can be identified;

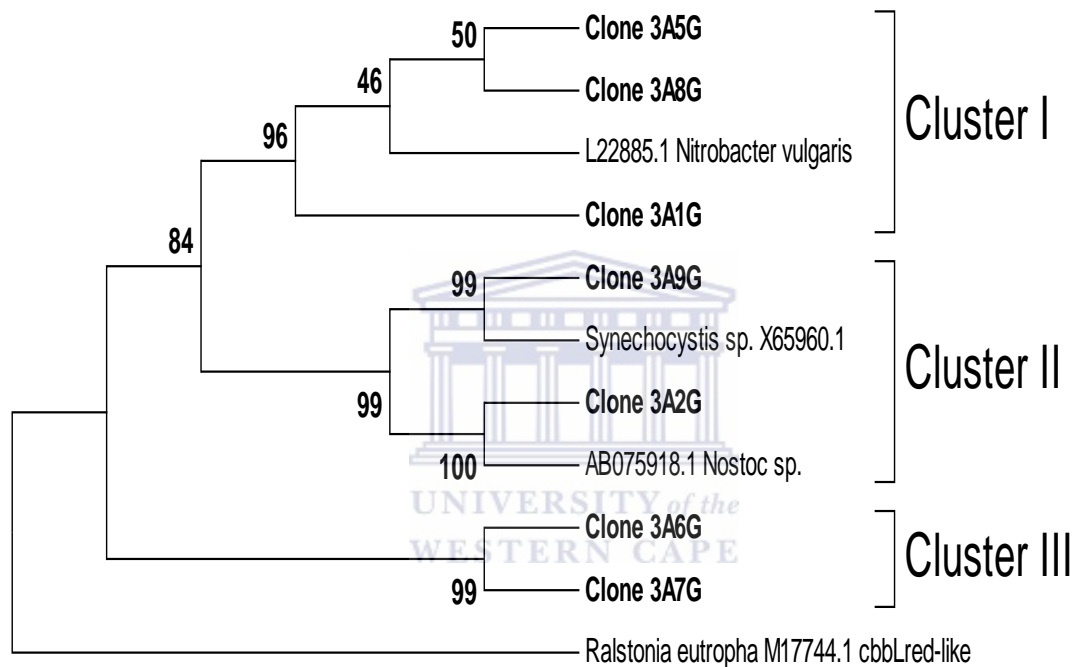


Figure 31: The evolutionary history was inferred for *cbbL* green-like sequences using the Neighbor-Joining method. The optimal tree with the sum of branch length = 2.78597539 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons (Pairwise deletion option). There were a total of 784 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4.

Cluster I:

The clones belonging to this cluster **3A5G**, **3A8G**, and **3A1G** were closely related to the *cbbL* green-like sequence of *Nitrobacter vulgaris* (L22885.1), which is an alpha-Proteobacteria (Appendix 3). Clones 3A5G and 3A8G cluster together with *Nitrobacter vulgaris*, forming a secondary cluster with Clone 3A1G being further in distance on the phylogenetic tree.

Cluster II:

Representative clones (**3A9G** and **3A2G**) are members of this cluster. Both clones have sequence similarity to the *cbbL* green-like gene of *Synechocystis* sp. (X65960.1), a member of the *Chroococcales* cyanobacteria. Clone **3A9G** forms a cluster 99% with *Synechocystis* sp., forming a sister clade. There is 100% similarity in the nucleotide sequence of *Nostoc* sp. (AB075918.1) and *cbbL* green-like Clone **3A9G**.

Cluster III:

Clones **3A6G** and **3A7G** formed the third cluster and established a position separately. The two clones formed a separate from the *cbbL* green-like sequences of previously cultured organisms. The phylogenetic tree generated indicates that the two clones cluster together 99% of the time. According to BLAST_n results the two clones are similar to *Nostoc* sp. (AB075918.1) however the percentage similarity (89%) may explain why the two clones form a cluster separate from *cbbL* green-like genes related to *Nostoc* sp. (AB075918.1).

4.4 Discussion

Previously, it has been reported that light-dependent CO₂ fixation by plants, algae and cyanobacteria is limited to the top few millimetres of the soil profile (Selesi *et al.*, 2005). However, no study has previously focused on light-dependent CO₂ fixation in oligotrophic Antarctic environments. This is the first reported investigation of the molecular diversity of RubisCO form I large subunit *cbbL* genes in Antarctic terrestrial soils. Evidence is presented of potential autotrophy by a range of previously uncultured *cbbL* harbouring microorganisms through the use of the functional gene coding for ribulose-1, 5-bisphosphate carboxylase oxygenase.

Only a few soil bacteria, including some heterotrophic and chemolithoautotrophic nitrifying bacteria, have been known to harbour *cbbL* genes (Selesi *et al.*, 2005; Tabita *et al.*, 2008a). Using previously designed primer pairs we showed that Antarctic hypolithic communities possess an abundance of *cbbL* related genes from a diverse range of cultured microorganisms.

Evidence of *cbbL* green-like phylotypes closely related to cyanobacteria and proteobacteria was presented. Two of the three main clusters obtained showed high bootstrap correlations to *cbbL* green-like sequences from both groups. Both groups, proteobacteria and cyanobacteria, have been previously detected through 16S rRNA gene analysis in Antarctic hypolithic communities (Khan, 2008; Pointing *et al.*, 2009). In contrast, the *cbbL* red-like genes obtained were relatively highly diverse. BLAST_n results indicate a strong similarity with previously uncultured taxa with various Proteobacterial

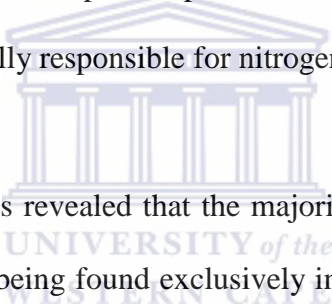
clades. Similarity to taxa known to carry out carbon fixation may possibly imply the presence of autotrophic groups based on the *cbbL* red-like gene. However, whether these genes are undergoing active transcription remains speculative.

Phylogeny based on *cbbL* sequence data is incongruent with the phylogeny based on 16S rRNA gene (Delwiche & Palmer, 1996). This may explain why certain clades in the phylogenetic tree contain sequences which belong to α -, and β - Proteobacteria (Appendix 3). This finding is consistent with previous studies using *cbbL* genes as phylogenetic markers. It is postulated that extensive horizontal gene transfer and gene duplication are responsible for the green-like and red-like split. However, in certain instances *cbbL* phylogeny is consistent with 16S rRNA gene based phylogeny. This suggests the need to supplement information obtained from *cbbL* libraries with 16S rRNA gene sequence data, in order to fully understand the structure of the total community.

Overall, the diversity of *cbbL* genes found in this study suggests the presence of microorganisms capable of carbon fixation in terrestrial Antarctic soils which is concurrent with previous findings (Aislabie *et al.*, 2006; Khan, 2008).

Chapter 5: Conclusions and Future Prospects

Examinations of the evolutionary relationships of specific functional bacterial groups have been performed before in several instances (Dedysh *et al.*, 2004; Leigh & Dodsworth, 2007; Selesi *et al.*, 2005). However, previous studies have explored microbial diversity in the Dry Valleys through 16S rRNA gene based techniques. This study explored, for the first time, members of microbial communities potentially involved in nutrient cycling. Key anabolic genes (*nifH* and *cbbL*) were successfully used as proxies to investigate members implicit in the particular nutrient cycles (nitrogen and carbon fixation respectively). Gene specific primers were essential in identification of microbial communities potentially responsible for nitrogen and carbon fixation.



16S rRNA gene DGGE analysis revealed that the majority of phylotypes were common in all samples with only a few being found exclusively in hypolithons. However, DGGE banding patterns suggests that hypolithons do support certain bacterial phylotypes that are not evident in open soil and non-translucent rock samples. DGGE analysis, as evidenced by statistical representation of banding patterns through MDS plots and similarity dendograms, indicates a lack of community variation, this is concurrent with other findings (Pointing *et al.*, 2009). This lack of significant variation among hypolith samples may indicate high selective pressures due to the environmental stresses in the Dry Valleys. Sequencing of DGGE bands was not undertaken during this study however previous studies based on 16S rRNA gene identification, have found a dominance of cyanobacteria and proteobacteria phylotypes in hypolithons (Khan, 2008; Pointing *et al.*, 2009).

The presence of various *nifH* gene phylotypes suggests the existence of diverse nitrogen fixing microorganisms in hypolithic communities. It can therefore be proposed that these hypolithic communities may play a pivotal role in nitrogen input in Antarctic terrestrial ecosystems. This proposal could be substantiated with the use of assays to determine nitrogen fixation rates in hypolithic communities. Preliminary data have recently been obtained (Cowan, D. A., Sohm, J. A., Makhalanyane, T., Capone, D. G., Green, T. G. A. & Cary, S. C; manuscript in preparation) which demonstrate detectable levels of acetylene reduction (a proxy for dinitrogen fixation) in Dry Valley hypolithic samples. However, this must be supplemented with assays determining nitrogen fixation at each stage (Figure 8, Chapter 1). Multivariate data analysis could be applied to investigate the effect of physiochemical variables on functional hypolithic communities. Additionally, transcriptomic analysis would shed light on which members of these communities are actively transcribing. Together these methods could be used to understand dinitrogen fixation and nitrogen cycling in Antarctic terrestrial environments.

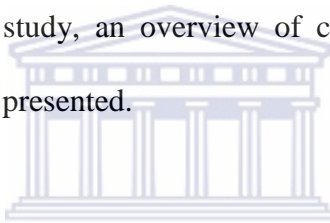
The *cbbL* gene was used to investigate carbon sequestration in hypolithic communities. The discovery of respective clusters of red-like and green-like *cbbL* genes indicates the presence of microorganisms which harbour these genes. The presence of these genes in soil, however, does not mean that active transcription is taking place. This merely implies the presence of communities previously and/or currently capable of carrying out carbon fixation.

Due to a relatively limited number of known *cbbL* sequences it is not possible to extrapolate any physiological or ecological function of RubisCO proteins in terrestrial environments because present data is based on *cbbL* gene diversity (Selesi *et al.*, 2005). Additionally currently available primers are based on available database sequences, which may bias the extent of diversity. Distantly related *cbbL* sequences (to those used for designing these primers) might not have been detected. Consequently, the diversity of green-like *cbbL* genes may be greater than what was discovered.

Investigation of the presence of specific phylotypes as indicated by presence of functional genes such as the *cbbL* gene may need to be accompanied by 16S rRNA gene community analysis. Although phylogeny based on the *nifH* gene is largely congruent with 16S rRNA gene phylogeny, this is clearly not the case with the *cbbL* gene.

Overall, this study presents a baseline from which perturbations in environmental conditions may be related to microbial communities, in particular putative C and N fixing microorganisms. The clone libraries depicted the presence of various phylotypes associated with known heterotrophic bacterial groups some of which have previously been observed in Antarctic environments (Aislabie *et al.*, 2006; Sjöling & Cowan, 2003). There is concurrency between the known groups found previously though 16S rRNA gene based studies and the current study based on functional gene diversity (Aislabie *et al.*, 2006; Khan, 2008; Smith *et al.*, 2006).

Knowledge of groups known to be involved in important ecosystem functioning processes such as C and N is essential. This information is important in developing hypothesis describing the influences of global change on these essential processes. Changes in microbial communities involved in C and N fixation, will contribute to a better understanding of biogeochemical cycling. Moreover, the study of low-diversity soil ecosystems such as cold deserts offer opportunities for elucidating relationships between biodiversity and ecosystem functioning. Changes in climate in the Antarctic provide an opportunity to assess the influences of temperature variations on the composition and functioning of soil communities, where soil food webs are simple relative to temperature (Barrett *et al.*, 2008). In this study, an overview of community members potentially important in these food webs is presented.



Future studies based on metatranscriptomic analysis of the functional genes investigated in this study are strongly recommended. mRNA retrieved from Antarctic communities will reveal functionally active communities harbouring both *nifH* and *cbbL* genes. These studies must incorporate 16S rRNA gene data, in order to design gene specific primers specific for phylotypes observed in Antarctic hypolithic communities.

It is also important, to obtain biogeochemical data indicating the presence of carbon and nitrogen in hypolithic communities. This data are important in possibly explaining the difference in *nifH* and *cbbL* OTU between hypolithic samples, open soil and non-translucent rock samples. This could further strengthen the link which has been suggested by some studies between gene diversity and ecosystem biogeochemistry (Liu *et al.*, 2003;

Selesi *et al.*, 2007). This will eventually contribute to understanding the ecological role of soil bacteria in C and N biogeochemical cycling.



Appendices


Appendix 1

Supplier	Location
Merck	Darmstadt, Germany
Bioline	England
Fermentas	Vilnius, Lithuania
Bio Rad	Munich, Germany
BDH,	England
Kimix	South Africa
Saarchem	South Africa
Promega	Madison, United States of America
Fluka	Germany
Sigma	Deisenhofen, Germany
<i>Taq</i> DNA polymerase	IMBM, South Africa
InqabaBiotech	South Africa
Integrated DNA Technologies (IDT)	South Africa
Difco (Becton Dickinson)	United States of America
Fermentas	Vilnius, Lithuania

Appendix 2

RFLP Pattern	Representative Clone	Accession number of Closest hit	Identity of sequences (%)	Highest BLAST_n match to cultured bacterium (C)	Accession number	Identity of sequences (%)	Phylogenetic affiliation based on BLAST_n search
1	Clone 2D11	AF049047.1	98%	<i>Spirirestis rafaensis</i> SRS70 dinitrogenase reductase (<i>nifH</i>) gene	DQ531685.1	93%	Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Spirirestis.
2	Clone A1	AF049044.1	98%	<i>Nostoc commune</i> SC clone 2 dinitrogenase	DQ531672.1	91%	Bacteria; Cyanobacteria; Nostocales;

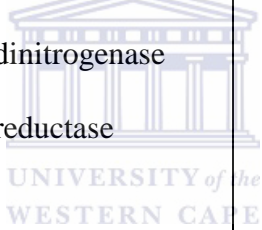
				reductase (nifH) gene,			Nostocaceae; Nostoc.
3	Clone A10	EU915052.1	94%	<i>Heliobacterium modesticaldum</i> Ice1, complete genome	CP000930.2	87%	Bacteria; Firmicutes; Clostridia; Clostridiales; Heliobacteriaceae; Heliobacterium.
4	Clone A11	AF049047.1	98%	<i>Spirirestis rafaelensis</i> SRS70 dinitrogenase reductase (nifH) gene	DQ531685.1	92%	Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Spirirestis.
5	Clone A2	AF049047.1	98%	<i>Scytonema sp.</i>	U73131.1	92%	Bacteria;

				dinitrogenase reductase (<i>nifH</i>) gene, partial cds			Cyanobacteria; Nostocales; Scytonemataceae; Scytonema
6	Clone A3			<i>Rhodococcus erythropolis</i> 	AP008957.1	100%	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
7	Clone A5	AF049045.1	99%	<i>Nostoc sp.</i> MFG-1 clone 1 dinitrogenase	DQ531687.1	95%	Bacteria; Cyanobacteria; Nostocales;

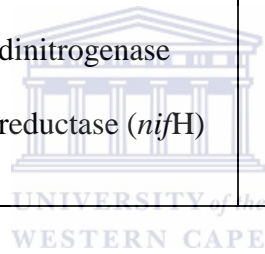
				reductase (<i>nifH</i>) gene			Nostocaceae; Nostoc.
8	Clone A6	AF049047.1	99%	<i>Spirirestis</i> <i>rafaelensis</i> SRS70 dinitrogenase reductase (<i>nifH</i>) gene, partial cds	DQ531685.1	93%	Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Spirirestis.
9	Clone A9	AF049047.1	99%	<i>Spirirestis</i> <i>rafaelensis</i> SRS70 dinitrogenase reductase (<i>nifH</i>) gene, partial cds	DQ531685.1	93%	Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Spirirestis.

10	Clone B10	AF049044.1	98%	<i>Tolypothrix</i> sp. LQ-10 dinitrogenase reductase (<i>nifH</i>) gene, partial cds	DQ531673.1	92%	Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Tolypothrix.
11	Clone B9	AF049047.1	99%	<i>Scytonema</i> sp. dinitrogenase reductase (<i>nifH</i>) gene, partial cds	U73131.1	92%	Bacteria; Cyanobacteria; Nostocales; Scytonemataceae; Scytonema.
12	Clone D11	GQ484838.1	97%	<i>Azotobacter</i> <i>vinelandii</i> DJ	CP001157.1	99%	Bacteria; Proteobacteria; Gammaproteobacteria;

							Pseudomonadales; Pseudomonadaceae; Azotobacter.
13	Clone D12	DQ142741.1	10%	<i>Azospirillum</i> <i>brasilense</i> strain AWC7 dinitrogenase reductase (<i>nifH</i>)	GQ161226.1	100%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales; Rhodospirillaceae; Azospirillum
14	Clone E9	AF049045.1	10%	<i>Nostoc sp. MFG-</i> 1 clone 1 dinitrogenase reductase (<i>nifH</i>) gene	DQ531687.1	95%	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.

15	Clone N1	AF049045.1	10%	<i>Nostoc sp.</i> MFG-1 clone 1 dinitrogenase reductase (<i>nifH</i>) gene	DQ531687.1	95%	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
16	Clone N2	AF049047.1	97%	<i>Scytonema sp.</i> dinitrogenase reductase 	U73131.1	92%	Bacteria; Cyanobacteria; Nostocales; Scytonemataceae; Scytonema
17	Clone N3	DQ776358.1	94%	<i>Polaromonas naphthalenivora</i> ns CJ2 (GENOME)	CP000529.1	91%	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

							Comamonadaceae; Polaromonas.
18	Clone N4	AF049047.1	99%	<i>Spirirestis</i> <i>rafaelensis</i> SRS70 dinitrogenase reductase (<i>nifH</i>)	DQ531685.1	92%	Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Spirirestis.



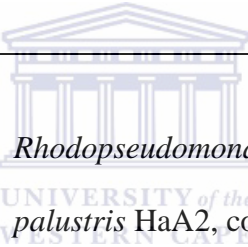
Appendix 3

RFLP Pattern	Representative Clone	Accession number of Closest Uncultured	Identity of sequences (%)	Highest BLAST _n match to cultured bacterium (C)	Accession number	Identity of sequences (%)	Phylogenetic affiliation based on BLAST _n search
1	Clone RA1	DQ149791.1	90%	<i>Xanthobacter autotrophicus</i> Py2, complete genome	CP000781.1	92%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Xanthobacteraceae; Xanthobacter.

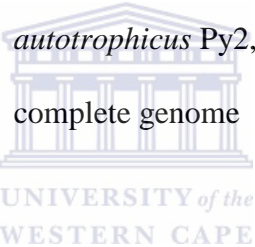
4	Clone RA2	DQ149754.1	87%	<i>Oscillochloris trichoides</i> strain C6 ribulose-1,5- biphosphate carboxylase/oxygenase large subunit form I (<i>cbbL</i>) gene, partial cds	DQ139401.1	85%	Bacteria; Chloroflexi; Chloroflexales; Oscillochloridaceae; Oscillochloris.
8	Clone RA3	EU450696.1	83%	<i>Variovorax paradoxus</i> S110 chromosome 1, complete sequence	CP001635.1	96%	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Comamonadaceae; Variovorax.


11	Clone RA4	AY572135.1	93%	<i>Bradyrhizobium</i> sp. BTAi1, complete genome	CP000494.1	93%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bradyrhizobium.
14	Clone RA5	EU450696.1	83%	<i>Variovorax paradoxus</i> S110 chromosome 1, complete sequence	CP001635.1	96%	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Comamonadaceae; Variovorax


20	Clone RA6	EU450746.1	83%	<i>Variovorax paradoxus</i> S110 chromosome 1, complete sequence	CP001635.1	95%	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Comamonadaceae; Variovorax
17	Clone RB1	AY572135.1	91%	<i>Bradyrhizobium sp.</i> ORS278,complete sequence	CP000494.1	93%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bradyrhizobium.


19	Clone RB2	EU450703.1	84%	<i>Rhodopseudomonas palustris</i> HaA2, complete genome	CP000250.1	90%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Rhodopseudomonas.
6	Clone RB4	EU450703.1	84%	 <i>Rhodopseudomonas palustris</i> HaA2, complete genome	CP000250.1	90%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Rhodopseudomonas
10	Clone RB5	AY572131.1	87%	<i>Thiohalomonas denitrificans</i> strain	GQ888591.1	82%	Bacteria; Proteobacteria;

				HLD3 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (<i>cbbL</i>) gene, partial cds			Gammaproteobacteria; Thiohalomonas.
16	Clone R1			<i>Schizaea pectinata</i> chloroplast partial rbcL gene for ribulose-1,5	AJ303409.1	80%	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Moniliformopses; Filicophyta; Filicopsida; Filicales; Schizaeaceae;

							Schizaea.
2	Clone AR	DQ149754.1	87%	<i>Xanthobacter</i>  <i>autotrophicus</i> Py2, complete genome	CP000781.2	85%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Xanthobacteraceae; Xanthobacter.
18	Clone RB1F	AY422925.1	81%	<i>Bradyrhizobium</i> sp. ORS278,complete sequence	CP000494.1	93%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

							Bradyrhizobiaceae; Bradyrhizobium.
5	Clone BR	AY572135.1	92%	<i>Bradyrhizobium sp.</i> BTai1 	CP000494.1	93%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bradyrhizobium.
9	Clone CR	AY572135.1	93%	<i>Bradyrhizobium sp.</i> BTai1,	CP000494.1	93%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae;

							Bradyrhizobium.
12	Clone D2R	AY572135.1	93%	<i>Bradyrhizobium sp.</i> ORS278,complete sequence 	CU234118.1	93%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bradyrhizobium.
3	Clone DR	AY422925.1	86%	<i>Bradyrhizobium sp.</i> BTAi1, complete genome	CP000494.1	93%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae;

							Bradyrhizobium.
13	Clone E1R	AY572131.1	93%	<i>Rhodopseudomonas palustris</i> TIE-1, complete genome 	CP001096.1	92%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Rhodopseudomonas.
7	Clone ER	AY572131.1	91%	<i>Paracoccus</i> sp. MU2A-22 ribulose-1,5-bisphosphate carboxylase/oxygenase	GQ452287.1	92%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae;

							Paracoccus.
15	Clone FR	AY422925.1	88%	<i>Bradyrhizobium sp.</i> BTai1, complete genome	CP000494.1	93%	



Appendix 4

RFLP Pattern	Representative Clone	Accession number of Closest hit Uncultured	Identity of sequences (%)	Highest BLAST_n match to cultured bacterium (C)	Accession number	Identity of sequences (%)	Phylogenetic affiliation based on BLAST_n search
1	Clone 3A1G	AY572174.1	100%	<i>Nitrobacter vulgaris</i>	L22885.1	98%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Nitrobacter
2	Clone 3A8G	AY572180.1	99%	<i>Nitrobacter vulgaris</i>	L22885.1	96%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

							Bradyrhizobiaceae; Nitrobacter
3	Clone 3A6G	AP009385.1	94%	<i>Nostoc sp.</i> PCC 7906 rbcL gene for ribulose-1,5- biphosphate carboxyla	AB075918.2	99%	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
4	Clone 3A9G			<i>Synechocystis sp.</i> UNIVERSITY of the WESTERN CAPE		98%	Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
5	Clone 3A2G	AP009385.1	98%	<i>Nostoc sp.</i> PCC 7906 rbcL gene for ribulose-1,5- biphosphate	AB075918.2	98%	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.

				carboxyla			
6	Clone 3A7G	AP009385.1	94%	<i>Nostoc sp. PCC</i> 7906 rbcL gene for ribulose-1,5- biphosphate carboxyla	AB075918.2	89%	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.



References

Aislabie, J. M., Chhour, K.-L., Saul, D. J., Miyauchi, S., Ayton, J., Paetzold, R. F. & Balks, M. R. (2006). Dominant bacteria in soils of Marble Point and Wright Valley, Victoria Land, Antarctica. *Soil Biology and Biochemistry* **38**, 3041-3056.

Andersen, D. T., McKay, C. P., R.A. Wharton, J. r. & Rummel, J. D. (1990). An Antarctic Research Outpost As A Model for Planetary Exploration. *Journal of The British Interplanetary Society* **43**, 499-504.

Andrew, W. D., Janet, L. G., Anita, P. J. & Tabita, F. R. (2005). Residues that influence *in vivo* and *in vitro* CbbR function in *Rhodobacter sphaeroides* and identification of a specific region critical for co-inducer recognition. *Molecular Microbiology* **57**, 1397-1414.

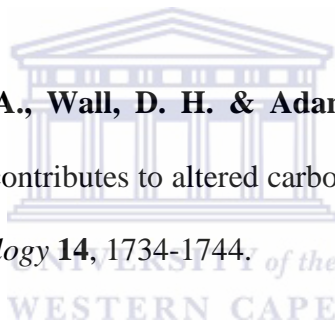
Ascaso, C. & Wierzchos, J. (2002). New approaches to the study of Antarctic lithobiontic microorganisms and their inorganic traces, and their application in the detection of life in Martian rocks. *International Microbiology* **5**, 215-222.

Atomi, H. (2002). Microbial enzymes involved in carbon dioxide fixation. *Journal of Bioscience and Bioengineering* **94**, 497-505.

Babalola, O. O., Kirby, B. M., Roes-Hill, M. L., Cook, A. E., Cary, S. C., Burton, S., G. & Cowan, D. A. (2009). Phylogenetic analysis of actinobacterial populations associated with Antarctic Dry Valley mineral soils. *Environmental Microbiology* **11**, 566-576.

Barnard, R., Roux, X. L., Hungate, B. A., Cleland, E. E., Blankinship, J. C., Barthes, L. & Leadley, P. W. (2006). Several components of global change alter nitrifying and denitrifying activities in an annual grassland. *Functional Ecology* **20**, 557-564.

Barrett, J. E., Virginia, R. A., Wall, D. H. & Adams, B. J. (2008). Decline in a dominant invertebrate species contributes to altered carbon cycling in a low-diversity soil ecosystem. *Global Change Biology* **14**, 1734-1744.



Berman-Frank, I., Lundgren, P. & Falkowski, P. (2003). Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria. *Research in Microbiology* **154**, 157-164.

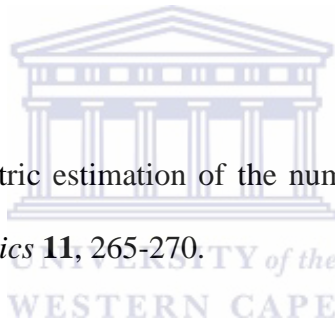
Bliss, L. C. & Gold, W. G. (1994). The patterning of plant communities and endaphic factors along the high Arctic coastline-implications for succession. *Canadian Journal of Botany*, 161 - 171.

Boyd, W. L. & Boyd, J. W. (1963). Soil Microorganisms of the McMurdo Sound Area, Antarctica. *Applied and Environmental Microbiology* **11**, 116-121.

Burgmann, H., Widmer, F., Sigler, W. V. & Zeyer, J. (2003). mRNA Extraction and Reverse Transcription-PCR Protocol for Detection of *nifH* Gene Expression by *Azotobacter vinelandii* in Soil. *Applied and Environmental Microbiology* **69**, 1928-1935.

Cameron, R. E. (1971). Antarctic soil microbial and ecological investigations. In *Research in the Antarctic*,. Edited by L. Quam & H. Porter: American Association for the Advancement of Science.

Chao, A. (1984). Non-parametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics* **11**, 265-270.



Chao, A., Ma, M. C. & Yang, M. C. K. (1993). Stopping rules and estimation for recapture debugging with unequal failure rates. *Biometrika* **80**, 193 - 201.

Chien, Y. T. & Zinder, S. H. (1996). Cloning, functional organization, transcript studies, and phylogenetic analysis of the complete nitrogenase structural genes (*nifHDK2*) and associated genes in the archaeon *Methanosarcina barkeri* 227. *Journal of Bacteriology* **178**, 143-148.

Church, M. J., Short, C. M., Jenkins, B. D., Karl, D. M. & Zehr, J. P. (2005). Temporal Patterns of Nitrogenase Gene (*nifH*) Expression in the Oligotrophic North Pacific Ocean. *Applied and Environmental Microbiology* **71**, 5362-5370.

Clegg, M. T. (1993). Chloroplast gene sequences and the study of plant evolution. *Proceedings of the National Academy of Sciences* **90**, 363-367.

Cockell, C. S. & Stokes, M. D. (2004). Ecology: Widespread colonization by polar hypoliths. *Nature* **431**, 414-414.

Coelho, M. R. R., Marriel, I. E., Jenkins, S. N., Lanyon, C. V., Seldin, L. & O'Donnell, A. G. (2009). Molecular detection and quantification of *nifH* gene sequences in the rhizosphere of sorghum (*Sorghum bicolor*) sown with two levels of nitrogen fertilizer. *Applied Soil Ecology* **42**, 48-53.

Committee on Antarctic Policy and Science, Polar Research Board, Commission on Geosciences Environment and Resources & National Research Council (1993). *Science and Stewardship in the Antarctic*. Washington, D.C: National Academy Press.

Committee on Principles of Environmental Stewardship for the Exploration and Study of Subglacial Environments & National Research Council (2007). *Exploration of Antarctic Subglacial Aquatic Environments: Environmental and Scientific Stewardship*. Washington, D.C.: The National Academy Press.

Convey, P. (1996). The influence of environmental characteristics on life history attributes of Antarctic terrestrial biota. *Biological Reviews* **71**, 191-225.

Convey, P. (1997). How are the life history strategies of Antarctic terrestrial invertebrates influenced by extreme environmental conditions? *Journal of Thermal Biology Special Issue to Commemorate the Life and Work of Professor Jack Grainger* **22**, 429-440.

Convey, P. & Stevens, M. I. (2007). Antarctic biodiversity. *Science* **317**, 1877 - 1878.

Cowan, D., Russell, N., Mamais, A. & Sheppard, D. (2002). Antarctic Dry Valley mineral soils contain unexpectedly high levels of microbial biomass. *Extremophiles* **6**, 431-436.

Cowan, D. A. & Ah Tow, L. (2004). Endangered Antarctic Environments. *Annual Review of Microbiology* **58**, 649-690.

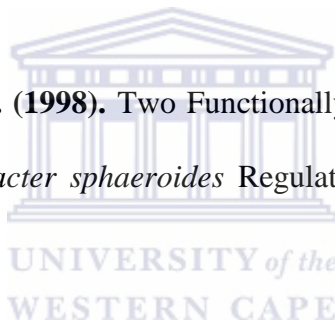
de la Torre, J. R., Goebel, B. M., Friedmann, E. I. & Pace, N. R. (2003). Microbial Diversity of Cryptoendolithic Communities from the McMurdo Dry Valleys, Antarctica. *Applied and Environmental Microbiology* **69**, 3858-3867.

Dedysh, S. N., Ricke, P. & Liesack, W. (2004). *nifH* and *nifD* phylogenies: an evolutionary basis for understanding nitrogen fixation capabilities of methanotrophic bacteria. *Microbiology* **150**, 1301-1313.

Delwiche, C. F. & Palmer, J. D. (1996). Rampant horizontal transfer and duplication of rubisco genes in eubacteria and plastids. *Molecular Biology and Evolution* **13**, 873-882.

Doran, P. T., Priscu, J. C., Lyons, W. B. & other authors (2002). Antarctic climate cooling and terrestrial ecosystem response. *Nature* **415**, 517-520.

Dubbs, J. M. & Tabita, F. R. (1998). Two Functionally Distinct Regions Upstream of the *cbbL* Operon of *Rhodobacter sphaeroides* Regulate Gene Expression. *Journal of Bacteriology* **180**, 4903-4911.



Dubbs, J. M., Bird, T. H., Bauer, C. E. & Tabita, F. R. (2000). Interaction of CbbR and RegA* Transcription Regulators with the *Rhodobacter sphaeroides* *cbbI* Promoter-Operator Region. *Journal of Biological Chemistry* **275**, 19224-19230.

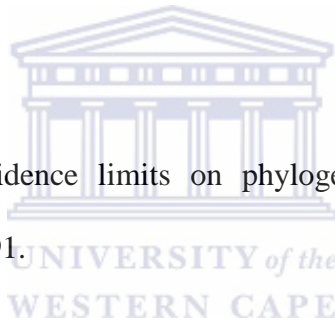
Elsaied, H. & Naganuma, T. (2001). Phylogenetic Diversity of Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase Large-Subunit Genes from Deep-Sea Microorganisms. *Applied and Environmental Microbiology* **67**, 1751-1765.

Falkowski, P. G. (1997). Evolution of the nitrogen cycle and its influence on the biological sequestration of CO₂ in the ocean. *Nature* **387**, 272-275.

Farrelly, V., Rainey, F. A. & Stackebrandt, E. (1995). Effect of genome size and rrn gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. *Applied and Environmental Microbiology* **61**, 2798-2801.

Fell, J. W., Scorzetti, G., Connell, L. & Craig, S. (2006). Biodiversity of micro-eukaryotes in Antarctic Dry Valley soils with <5% soil moisture. *Soil Biology and Biochemistry* **38**, 3107-3119.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**, 783-791.



Frey, J. C., Angert, E. R. & Pell, A. N. (2006). Assessment of biases associated with profiling simple, model communities using terminal-restriction fragment length polymorphism-based analyses. *Journal of Microbiological Methods* **67**, 9-19.

Gibson, J. L. & Tabita, F. R. (1997). Analysis of the *cbbXYZ* operon in *Rhodobacter sphaeroides*. *Journal of Bacteriology* **179**, 663-669.

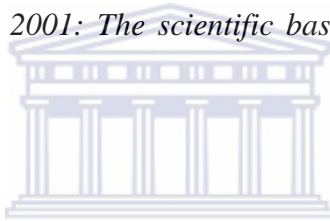
Griffiths, R. I., Whiteley, A. S., O'Donnell, A. G. & Bailey, M. J. (2000). Rapid Method for Coextraction of DNA and RNA from Natural Environments for Analysis of

Ribosomal DNA- and rRNA-Based Microbial Community Composition. *Applied and Environmental Microbiology* **66**, 5488-5491.

Hong, S., Bunge, J., Leslin, C., Jeon, S. & Epstein, S. S. (2009). Polymerase chain reaction primers miss half of rRNA microbial diversity. *ISME Journal*.

Horowitz, N. & Hubbard, R. C. J. (1972). Microbiology the Dry Valleys of Antarctica. *Science* **176**, 242-245.

IPCC (2001). *Climate change 2001: The scientific basis*. Cambridge, UK: Cambridge University Press.



IPCC (2007). Climate change 2007. In *Climate change impacts, adaptations and vulnerability Working Group II*. Edited by Intergovernment Panel on Climate Change. Geneva, Switzerland: IPCC.

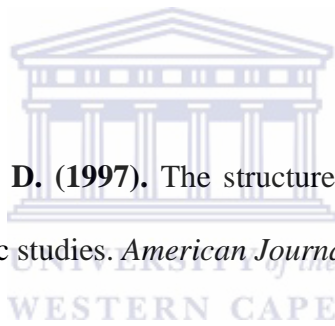
Jacka, T. H. & Budd, W. F. (1998). Detection of temperature and sea-ice-extent changes in the Antarctic and Southern Ocean 1949-96. *Annals of Glaciology* **27**, 553 - 559.

Jenkins, B. D., Steward, G. F., Short, S. M., Ward, B. B. & Zehr, J. P. (2004). Fingerprinting Diazotroph Communities in the Chesapeake Bay by Using a DNA Macroarray. *Applied and Environmental Microbiology* **70**, 1767-1776.

Jungblut, A. D. & Neilan, B. A. (2009). NifH gene diversity and expression in a microbial mat community on the McMurdo Ice Shelf, Antarctica. *Antarctic Science* **Forthcoming**, 1-6.

Kallersjo, M., Farris, J. S., Chase, M. W., Bremer, B., Fay, M. F., Humphries, C. J., Petersen, G., Seberg, O. & Bremer, K. (1998). Simultaneous parsimony jackknife analysis of 2538 *rbcL* DNA sequences reveals support for major clades of green plants, land plants, seed plants and flowering plants. *Plant Systematics and Evolution* **213**, 259 - 287.

Kellogg, E. A. & Juliano, N. D. (1997). The structure and function of RuBisCO and their implications for systematic studies. *American Journal of Botany* **84**, 413-.



Kemp, P. F. & Aller, J. Y. (2004). Estimating prokaryotic diversity: When are 16S rDNA libraries large enough? *Limnology and Oceanography: Methods* **2**, 114-125.

Keyster, M. (2007). Spatial distribution of cyanobacterial phylotypes in Antarctic Dry Valley soil biotopes. In *The Department of Biotechnology*, pp. 133. Cape Town: University of the Western Cape.

Khan, N. (2008). Characterisation of microbial communities associated with hypolithic environments in Antarctic Dry Valley soils. In *The Department of Biotechnology*. Cape Town: University of the Western Cape.

King, G. M., Kirchman, D., Salyers, A. A., Schlesinger, W. & Tiedje, J. M. (2006). Global Environmental Change: Microbial Contributions Microbial Solutions: American Society for Microbiology.

Kirshtein, J. D., Paerl, H. W. & Zehr, J. P. (1991). Amplification, Cloning, and Sequencing of a *nifH* Segment from Aquatic Microorganisms and Natural Communities. *Applied and Environmental Microbiology* **57**, 2645-2650.

Lal, R. (2008). Carbon sequestration. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**, 815-830.

Leigh, J. A. & Dodsworth, J. A. (2007). Nitrogen Regulation in Bacteria and Archaea. *Annual Review of Microbiology* **61**, 349-377.

Liu, X., Tiquia, S. M., Holguin, G. & other authors (2003). Molecular Diversity of Denitrifying Genes in Continental Margin Sediments within the Oxygen-Deficient Zone off the Pacific Coast of Mexico. *Applied and Environmental Microbiology* **69**, 3549-3560.

Luque, I. & Forchhammer, K. (2008). Nitrogen Assimilation and C/N Balance Sensing. In *The Cyanobacteria: Molecular Biology, Genomics and Evolution*, pp. 484. Edited by A. Herrero & E. Flores. Norfolk, UK: Caister Academic Press.

Mataloni, G., Tell, G. & Wynn-Williams, D. D. (2000). Structure and diversity of soil algal communities from Cierva Point (Antarctic Peninsula). *Polar Biology* **23**, 205-211.

Melody, B. B., Ross, A. V. & Diana, H. W. (2001). Organic carbon cycling in Taylor Valley, Antarctica: quantifying soil reservoirs and soil respiration. *Global Change Biology* **7**, 113-125.

Miller, D. N., Bryant, J. E., Madsen, E. L. & Ghiorse, W. C. (1999). Evaluation and Optimization of DNA Extraction and Purification Procedures for Soil and Sediment Samples. *Applied and Environmental Microbiology* **65**, 4715-4724.

Moorhead, D. L., Doran, P. T., Fountain, A. G., Lyons, W. B., McKnight, D. M., Priscu, J. C., Virginia, R. A. & Wall, D. H. (1999). Ecological Legacies: Impacts on Ecosystems of the McMurdo Dry Valleys. *BioScience* **49**, 1009-1019.

Muyzer, G., de Waal, E. C. & Uitterlinden, A. G. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology* **59**, 695-700.

Nakatsubo, T. & Ino, Y. (1986). Nitrogen cycling in an Antarctic ecosystem. 1. Biological nitrogen fixation in the vicinity of Syowa Station. *Memoirs of National Institute of Polar Research Ser E, Biology and medical science* **37**, 1-10.

Nannipieri, P., Ascher, J., Ceccherini, M. T., Landi, L., Pietramellara, G. & Renella, G. (2003). Microbial diversity and soil functions. *European Journal of Soil Science* **54**, 655-670.

Onofri, S., Selbmann, L., Zucconi, L. & Pagano, S. (2004). Antarctic microfungi as models for exobiology. *Planetary and Space Science Exploring Mars Surface and its Earth Analogues* **52**, 229-237.

Pearce, D. A., Bridge, P., D. , Hughes, K., A., Sattler, B., Psenner, R. & Russell, N., J. (2009). Microorganisms in the atmosphere over Antarctica. *FEMS Microbiology Ecology* **69**, 143-157.

Pointing, S. B., Chan, Y., Lacap, D. C., Lau, M. C. Y., Jurgens, J. A. & Farrell, R. L. (2009). Highly specialized microbial diversity in hyper-arid polar desert. *Proceedings of the National Academy of Sciences* **106**, 19964-19969.

Poly, F., Monrozier, L. J. & Bally, R. (2000). Improvement in the RFLP procedure for studying the diversity of *nifH* genes in communities of nitrogen fixers in soil. *Research in Microbiology* **152**, 95-103.

Poly, F., Ranjard, L., Nazaret, S., Gourbiere, F. & Monrozier, L. J. (2001). Comparison of *nifH* Gene Pools in Soils and Soil Microenvironments with Contrasting Properties. *Applied and Environmental Microbiology* **67**, 2255-2262.

Postgate, J. (1998). *Nitrogen Fixation*, 3rd edn. Cambridge UK: Cambridge University Press.

Rees, D. C., Akif Tezcan, F., Haynes, C. A., Walton, M. Y., Andrade, S., Einsle, O. & Howard, J. B. (2005). Structural basis of biological nitrogen fixation. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **363**, 971-984.

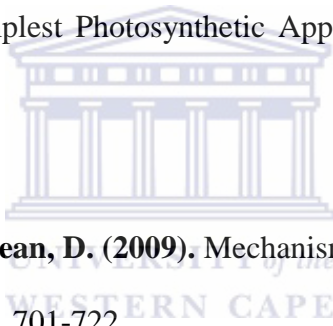
Rochelle, M. S., Susanna, A. W., Joseph, J. G., Ian, R. M. & Cary, S. C. (2009). Microbial biodiversity of thermophilic communities in hot mineral soils of Tramway Ridge, Mount Erebus, Antarctica. *Environmental Microbiology* **11**, 715-728.

Russelle, M. P. & Birr, A. S. (2004). Biological Nitrogen Fixation. Large-Scale Assessment of Symbiotic Dinitrogen Fixation by Crops: Soybean and Alfalfa in the Mississippi River Basin. *American Society of Agronomy* **96**, 1754-1760.

Saitou N & Nei M (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406-425.

Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: a laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press,.

Sattley, W. M., Madigan, M. T., Swingley, W. D. & other authors (2008). The Genome of *Heliobacterium modesticaldum*, a Phototrophic Representative of the Firmicutes Containing the Simplest Photosynthetic Apparatus. *Journal of Bacteriology* **190**, 4687- 4696.



Seefeldt, L., Hoffman, B. & Dean, D. (2009). Mechanism of Mo-dependent nitrogenase. *Annual Review in Biochemistry*, 701-722.

Selesi, D., Schmid, M. & Hartmann, A. (2005). Diversity of Green-Like and Red-Like Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase Large-Subunit Genes (*cbbL*) in Differently Managed Agricultural Soils. *Applied and Environmental Microbiology* **71**, 175-184.

Selesi, D., Pattis, I., Schmid, M., Kandeler, E. & Hartmann, A. (2007). Quantification of bacterial RuBisCO genes in soils by *cbbL* targeted real-time PCR. *Journal of Microbiological Methods* **69**, 497-503.

Shaw, M. R. & John, H. (2001). Response of nitrogen cycling to simulated climate change: differential responses along a subalpine ecotone. *Global Change Biology* **7**, 193-210.

Shively, J. M., Devore, W., Stratford, L., Porter, L., Medlin, L. & S. E. Stevens, J. (1986). Molecular evolution of the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO). *FEMS Microbiology Letters* **37**, 251-257.

Shively, J. M., English, R. S., Baker, S. H. & Cannon, G. C. (2001). Carbon cycling: the prokaryotic contribution. *Current Opinion in Microbiology* **4**, 301-306.

Shively, J. M., van Keulen, G. & Meijer, W. G. (2003). Something From Almost Nothing: Carbon Dioxide Fixation in Chemoautotrophs. *Annual Review of Microbiology* **52**, 191-230.

Simonet, P., Grosjean, M. C., Misra, A. K., Nazaret, S., Cournoyer, B. & Normand, P. (1991). Frankia genus-specific characterization by polymerase chain reaction. *Applied and Environmental Microbiology* **57**, 3278-3286.

Sinigalliano, C. D., Jones, R. D., Kuhn, D. N. & Guerrero, M. A. (2003). Visualization of the *cbbL* Gene in Healthy and Starved Chemoautotrophic Nitrifying Bacteria. *Current Microbiology* **47**, 0077-0083.

Sjöling, S. & Cowan, D. A. (2003). High 16S rDNA bacterial diversity in glacial meltwater lake sediment, Bratina Island, Antarctica. *Extremophiles* **7**, 275-282.

Smith, J., Tow, L., Stafford, W., Cary, C. & Cowan, D. (2006). Bacterial Diversity in Three Different Antarctic Cold Desert Mineral Soils. *Microbial Ecology* **51**, 413-421.

Smith, K. A., Ball, T., Conen, F., Dobbie, K. E., Massheder, J. & Rey, A. (2003). Exchange of greenhouse gases between soil and atmosphere: interactions of soil physical factors and biological processes. *European Journal of Soil Science* **54**, 779-791.

Stevens, M. I. & Hogg, I. D. (2002). Expanded distributional records of Collembola and Acari in southern Victoria Land, Antarctica. *Pedobiologia* **46**, 485-495.

Steward, G. F., Zehr, J. P., Jellison, R., Montoya, J. P. & Hollibaugh, J. T. (2004). Vertical Distribution of Nitrogen-Fixing Phylotypes in a Meromictic, Hypersaline Lake. *Microbial Ecology* **47**, 30-40.

Tabita, F. R., Janet, L. G., Botho, B., Lubbert, D. & Wim, G. M. (1992). Uniform designation for genes of the Calvin-Benson-Bassham reductive pentose phosphate pathway of bacteria. *FEMS Microbiology Letters* **99**, 107-110.

Tabita, F. R. (1999). Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: A different perspective. *Photosynthesis Research* **60**, 1-28.

Tabita, F. R., Hanson, T. E., Li, H., Satagopan, S., Singh, J. & Chan, S. (2007). Function, Structure, and Evolution of the RuBisCO-Like Proteins and Their RubisCO Homologs. *Microbiology and Molecular Biology Reviews* **71**, 576-599.

Tabita, F. R., Hanson, T. E., Satagopan, S., Witte, B. H. & Kreel, N. E. (2008a). Phylogenetic and evolutionary relationships of RubisCO and the RubisCO-like proteins and the functional lessons provided by diverse molecular forms. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**, 2629-2640.

Tabita, F. R., Satagopan, S., Hanson, T. E., Kreel, N. E. & Scott, S. S. (2008b). Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *Journal of Experimental Botany* **59**, 1515-1524.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. . *Molecular Biology and Evolution* **24**, 1596-1599.

Thiel, T. (1993). Characterization of genes for an alternative nitrogenase in the cyanobacterium *Anabaena variabilis*. *Journal of Bacteriology* **175**, 6276-6286.

Thornton, P. E., Doney, S. C., Lindsay, K. & other authors (2009). Carbon-nitrogen interactions regulate climate-carbon cycle feedbacks: results from an atmosphere-ocean general circulation model. *Biogeosciences Discuss* **6**, 3303-3354.

Tolli, J. & King, G. M. (2005). Diversity and Structure of Bacterial Chemolithotrophic Communities in Pine Forest and Agroecosystem Soils. *Applied and Environmental Microbiology* **71**, 8411-8418.

Torsvik, V. & Øvreås, L. (2002). Microbial diversity and function in soil: from genes to ecosystems. *Current Opinion in Microbiology* **5**, 240-245.

Tourova, T. P., Spiridonova, E. M., Berg, I. A., Kuznetsov, B. B. & Sorokin, D. Y. (2006). Occurrence, phylogeny and evolution of ribulose-1,5-bisphosphate carboxylase/oxygenase genes in obligately chemolithoautotrophic sulfur-oxidizing bacteria of the genera *Thiomicrospira* and *Thioalkalimicrobium*. *Microbiology* **152**, 2159-2169.

Treonis, A. M., Wall, D. H. & Virginia, R. A. (1999). Invertebrate Biodiversity in Antarctic Dry Valley Soils and Sediments. *Ecosystems* **2**, 482-492.

Turner, J., Colwell, S., R. , Marshall, G., J., Lachlan-Cope, T., A., Carleton, A., M. , Jones, P., D. , Lagun, V., Reid, P., A. & Iagovkina, S. (2005). Antarctic climate change during the last 50 years. *International Journal of Climatology* **25**, 279-294.

United Nations (2009). Framework Convention on Climate Change on World. In <http://www.unfccc.de>.

Videmšek, U., Hagn, A., Suhadolc, M., Radl, V., Knicker, H., Schloter, M. & Vodnik, D. (2009). Abundance and Diversity of CO₂-fixing Bacteria in Grassland Soils Close to Natural Carbon Dioxide Springs. *Microbial Ecology* **58**, 1-9.

Wang, S.-Z., Chen, J.-S. & Johnson, J. L. (1988). The presence of five *nifH*-like sequences in *Clostridium pasteurianum*: sequence divergence and transcription properties. *Nucleic Acids Research* **16**, 439-454.

Warren-Rhodes, K., Rhodes, K., Pointing, S., Ewing, S., Lacap, D., Gómez-Silva, B., Amundson, R., Friedmann, E. & McKay, C. (2006). Hypolithic Cyanobacteria, Dry Limit of Photosynthesis, and Microbial Ecology in the Hyperarid Atacama Desert. *Microbial Ecology* **52**, 389-398.

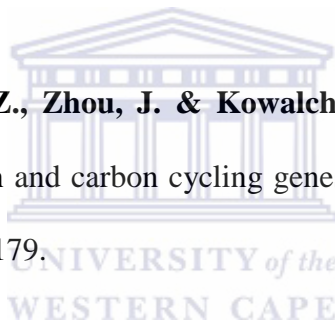
Weisburd, S. (1986). A Forest Grows in Antarctica. *Science News* **129**, 148.

Wynn-Williams, D. (1989). Ecological aspects of Antarctic microbiology. *Advanced Microbial Ecology* **11**, 71-76.

Wynn-Williams, D. D. (1988). Television Image Analysis of Microbial Communities in Antarctic Fellfields. *Polarforschung* **58**, 239-249.

Xu, H. H. & Tabita, F. R. (1996). Ribulose-1,5-bisphosphate carboxylase/oxygenase gene expression and diversity of Lake Erie planktonic microorganisms. *Applied and Environmental Microbiology* **62**, 1913-1921.

Yergeau, E., Kang, S., He, Z., Zhou, J. & Kowalchuk, G. A. (2007a). Functional microarray analysis of nitrogen and carbon cycling genes across an Antarctic latitudinal transect. *ISME Journal* **1**, 163-179.



Yergeau, E., Newsham, K. K., Pearce, D. A. & Kowalchuk, G. A. (2007b). Patterns of bacterial diversity across a range of Antarctic terrestrial habitats. *Environmental Microbiology* **9**, 2670-2682.

Zakhia, F., Jungblut, A.-D., Taton, A., Vincent, W. F. & Wilmotte, A. (2008). Cyanobacteria in Cold Ecosystems. In *Psychrophiles: from Biodiversity to Biotechnology*, pp. 121-135.

Zeglin, L., Sinsabaugh, R., Barrett, J., Gooseff, M. & Takacs-Vesbach, C. (2009). Landscape Distribution of Microbial Activity in the McMurdo Dry Valleys: Linked Biotic Processes, Hydrology, and Geochemistry in a Cold Desert Ecosystem. *Ecosystems* **12**, 562-573.

Zehr, J., P., Jenkins, B., D., Short, Steven, M. & Steward, G., F. (2003). Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environmental Microbiology* **5**, 539-554.

Zehr, J. P. & McReynolds, L. A. (1989). Use of Degenerate Oligonucleotides for Amplification of the *nifH* Gene from the Marine Cyanobacterium *Trichodesmium thiebautii*. *Applied and Environmental Microbiology* **55**, 2522-2526.

Zehr, J. P., Mellon, M. T. & Hiorns, W. D. (1997). Phylogeny of cyanobacterial *nifH* genes: evolutionary implications and potential applications to natural assemblages. *Microbiology* **143**, 1443-1450.

Zehr, J. P., Waterbury, J. B., Turner, P. J., Montoya, J. P., Omoregie, E., Steward, G. F., Hansen, A. & Karl, D. M. (2001). Unicellular cyanobacteria fix N₂ in the subtropical North Pacific Ocean. *Nature* **412**, 635-638.

Zehr, J. P. & Ward, B. B. (2002). Nitrogen Cycling in the Ocean: New Perspectives on Processes and Paradigms. *Applied and Environmental Microbiology* **68**, 1015-1024.

Zhang, Y., Dong, J., Yang, Z., Zhang, S. & Wang, Y. (2008). Phylogenetic diversity of nitrogen-fixing bacteria in mangrove sediments assessed by PCR–denaturing gradient gel electrophoresis. *Archives of Microbiology* **190**, 19-28.

