# Spatial distribution of cyanobacterial phylotypes in Antarctic Dry Valley soil biotopes

***Marshall Keyster***


A thesis submitted in partial fulfillment of the requirements for the degree of

MAGISTER SCIENTIAE (M.Sc)

Department of Biotechnology,

University of the Western Cape

Bellville

Supervisor: Prof. D.A. Cowan

**November 2007**

## ACKNOWLEDGEMENTS

My heavenly Father for providing me guidance, strength and wisdom

I am extremely grateful to my supervisor Professor Don Cowan for providing the opportunity and necessary guidance to complete my project

To my co-supervisor Dr. Lukas Rohr for his extreme patience, assistance, technical support and structuring of the thesis

My Parents, for their love, support and understanding through all my years of study

A special thanks to Elize and Deon for providing me with a quiet home to finish the thesis

Thank you to Jessica, for her steadfast patience and support through the difficult times

A special thanks to my friends and colleagues in U.W.C Department of Biotechnology especially Morné, Ashwil, Joseph, Billi, Babelle, Parik, Anastasia, Dean, Pauline, Walter, Wycleffe, Kazi and Andrew N

Everyone in IMBM, especially Dr. Heidi Goodman and the technical staff for endless assistance. To the rest of the group thank you very much for humor, advice and moral support, it has been a rough ride but one I would never forget

The National Research Foundation for providing the necessary funding

*"A good theory is characterized by the fact that it makes a number of predictions that could in principle be disproved or falsified by observation. Each time new experiments are observed to agree with the predictions the theory survives, and our confidence in it is increased; but if ever a new observation is found to disagree, we have to abandon or modify the theory. At least that is what is supposed to happen, but you can always question the competence of the person who carried out the observation".*



**-Karl Popper** (Philosopher of science) as phrased by Stephen Hawking in the book A Brief History of Time (pp. 11)

***This thesis is dedicated to my late father Adam, brother André and cousin Jody***

## ABSTRACT

### Spatial distribution of cyanobacterial phylotypes in Antarctic Dry Valley soil biotopes

**Marshall Keyster**

M.Sc thesis, Department Biotechnology,
University of the Western Cape

In the past studies, on cyanobacterial diversity were mainly done using microscopic methods based on differences in morphology. However, recent advances in molecular methods have enabled the analysis of cyanobacterial diversity using PCR-based approaches. Snap-shot techniques such as 16S rDNA library construction and DGGE have allowed for increased access and better understanding of cyanobacterial diversity from diverse biotopes. These environments include sea water, deep-sea regions, deserts, forest soils and temperate soils. Due to the wide distribution of Cyanobacteria, increasing interest has been shown in the Antarctic continent. Cyanobacteria have been proposed as the dominant bacterial phylum in the extreme Antarctic ecosystem. However, studies on cyanobacterial diversity have primarily focused on the Antarctic aquatic systems (lakes and ponds), leaving a gap in the knowledge on terrestrial cyanobacteria from Antarctica. In this investigation, 16S rDNA analysis and DGGE profiling were used to study the diversity of cyanobacteria along a transect of increasing altitude in the Miers Valley, Eastern Antarctica.

During the first part of this study high molecular weight DNA was extracted from eleven soil samples along the transect. The DNA was of high quality in terms of purity and could be used directly in subsequent PCR amplifications without further purification. A nested-PCR strategy was employed using a cyanobacterial-specific forward primer combined with a universal reverse primer in the first round followed by a second round amplification targeting the cyanobacterial 16S rDNA more specifically with cyanobacterial specific primers. Following the successful PCR amplification using the extracted metagenomic DNA the amplicons were used in DGGE experiments for the analysis of the cyanobacterial community along the transect. A homogeneity test revealed the all samples were highly homogenous in terms of cyanobacterial community structure. This allowed for the selection of only the homogeneous sample (J) further analysis. An analysis of DGGE profiles suggested that altitude, spatial separation and relative soil humidity may play key roles in shaping the cyanobacterial community structure. DGGE bands were present in all samples however, the profile also indicated variations in cyanobacterial diversity and that some phylotypes are unique to specific samples.



In the second part of the study, putative cyanobacterial 16S rDNA libraries were constructed from four samples at 171 m, 184 m, 384 m and 634 m a.s.l and 96 clones of each library were subjected to ARDRA analysis. In contrast to the outcome of the DGGE analysis there was very little overlap of ARDRA groups between the four libraries. Numbers of observed ARDRA groups were used to

estimate the total number of ARDRA groups present in the libraries based on the $S_{chao1}$ richness estimator, which suggested unrealistically high cyanobacterial diversity, with an estimated total of 332 ARDRA groups in the four libraries.

A total of 101 clones were selected from the libraries for sequencing, and BLAST results revealed no cyanobacterial sequences from samples 384 and 634 while a high percentage of cyanobacterial sequences were identified among the clones of the lower altitude samples (70% and 55% for 171 and 184, respectively). *Anabaena* (25%), *Chroococcidiopsis* (17%), *Cylindrospermum* (54%) and *Oscillatoria* (4%) were the four cyanobacterial species identified with very low confidence after sequence classification. Acidobacteria followed by Actinobacteria dominated the two samples 384 and 634. Interestingly, sequences classified and grouped as unclassified (5 clones) were also present only in samples 384 and 634, indicating high numbers of novel species in these samples. The sequencing results showed overlapping of sequences between samples similar to DGGE profiles but contrasted with the ARDRA profile which showed very high diversity and numbers.

## DECLARATION

I declare that *Spatial distribution of cyanobacterial phylotypes in Antarctic Dry Valley soil biotopes* 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.

**Marshall Keyster** **November 2007**

---

## TABLE OF CONTENTS

## CHAPTER TWO: MATERIALS AND METHODS

## LIST OF ABBREVIATIONS

| | |
|---|---|
| ARDRA | Amplified Ribosomal Restriction Analysis |
| APS | Ammonium peroxidisulphate |
| a.s.l | above sea level |
| ATP | adenosine triphosphate |
| BLAST | Basic Local Alignment Search Tool |
| bp | base pair |
| cm | centimeter |
| DGGE | Denaturing Gradient Gel Electrophoresis |
| dNTPs | deoxyribonucleic-5'-triphosphate |
| DNA | deoxyribonucleic acid |
| °C | degrees Celsius |
| *E. coli* | *Escherichia coli* |
| EtOH | ethanol |
| EDTA | ethylene diamine tetraacetic acid |
| Fig. | figure |
| g | gram |
| GC-clamp | Guanosine-cytosine clamp |
| × g | centrifugal force |
| μg | microgram |
| μg $ml^{-1}$ | microgram per milliliter |
| $H_2SO_4$ | Sulfuric acid |
| IPTG | isopropyl β-D-thiogalactosidase |
| kb | kilo basepairs |
| kV | kilo Volt |
| l | liter |
| LB | Luria-Bertani |
| µl | micro liter |
| µF | micro Farad |

| | |
|---|---|
| M | molar |
| ml | milliliter |
| min | minute |
| mM | millimolar |
| NaCl | sodium chloride |
| NaOH | sodium hydroxide |
| ng | nano gram |
| $NH_4AOc$ | Ammonium acetate |
| Ω | Ohm |
| OD | optical density |
| PCR | Polymerase Chain Reaction |
| rpm | revolutions per minute |
| rDNA | ribosomal deoxyribonucleic acid |
| RDP | Ribosomal Database Project |
| RH | Relative humidity |
| rRNA | Ribosomal ribonucleic acid |
| s | second |
| SDS | sodium dodecyl sulphate |
| TAE | Tris acetic acid EDTA |
| TB | Terrific Broth |
| TE | Tris EDTA |
| TEMED | Tetramethylethyldiamine |
| U | unit |
| UHQ | ultra high quality |
| UV | ultra violet |
| V | volts |
| VBNC | Viable but non-culturable |
| v/v | volume per volume |
| w/v | weight per volume |
| X-gal | 5-bromo-4-chloro-3-indolyl-β-D-galactoside |

## LIST OF FIGURES

### CHAPTER 1



### CHAPTER 3

## CHAPTER 4

## LIST OF TABLES

### CHAPTER 1



### CHAPTER 2





### CHAPTER 3

# CHAPTER ONE

## Literature Review

### 1.1 Introduction

Soil microbial diversity in the natural environment is widely acknowledged to be severely underestimated (Fierer and Jackson, 2006). Initial studies on microbial diversity from soils involved culturing methods without realizing the bias introduced by the selective culturing conditions. The identification of bacteria from various soils enables microbial ecologists to group these microorganisms into different operational taxonomic entities thereby defining which types are present (composition), the total number of types (richness) and the relative abundance of each type (structure) (Gordon *et al.*, 2000).

However, limitations related to standard plate isolation methods existed. No specific culturing conditions that are required to successfully isolate the majority of bacteria exists (Janssen *et al*., 2002). Another restriction of culture-dependent strategies is the viable but not culturable (VBNC) state of some microorganisms. The VBNC state is believed to be either a process of self-preservation whereby the bacterium ensures survival during environmental stress, or an end-of-life-cycle process (McDougald *et al*., 1998).

Torsvik et al. (1990) investigated the reassociation kinetics of isolated microbial community DNA and estimated the bacterial diversity to be $10^3$ different genomes per sample. This figure was found to be 200 times larger than that obtained from culturing methods. This result indicated that standard plating techniques only

accessed a minute fraction of the soil microbial flora. Recent investigations of reassociation kinetics of soil microbial flora from environmental samples indicated that the diversity was underestimated a thousand fold (Gans *et al*., 2005).

Most of the Earth's prokaryotes are found in the open ocean and in soils. The numbers of these cells are believed to be in the order of $10^{29}$ to $10^{30}$ (Whitman *et al*., 1998). However, limitations of collecting the total number of prokaryotes on Earth exist. These include, limited sampling of some environments followed by correct molecular protocols to asses the complete diversity that exist. This means that the overall level of endemicity on Earth is still undetermined.

The high degrees of novelty greatly expand the Earth's microbial diversity making it impossible to calculate and isolate the total number of prokaryotes that exist (Bull and Stach, 2004). This give rise to the curiosity of finding novel microorganisms from previous unsampled environments. Some of these environments are called extreme environments where no complex eukaryotes survive (Tindall, 2004). These environments include regions of high temperature, low temperature, high salinity, high pressure, high acidity, nutrient deficiency, and low water content.

These environments have gained considerable scientific attention in the last decade, not only for novel microorganism identification but also because of the unidentified byproducts of these microorganisms.

## 1.2 Antarctica

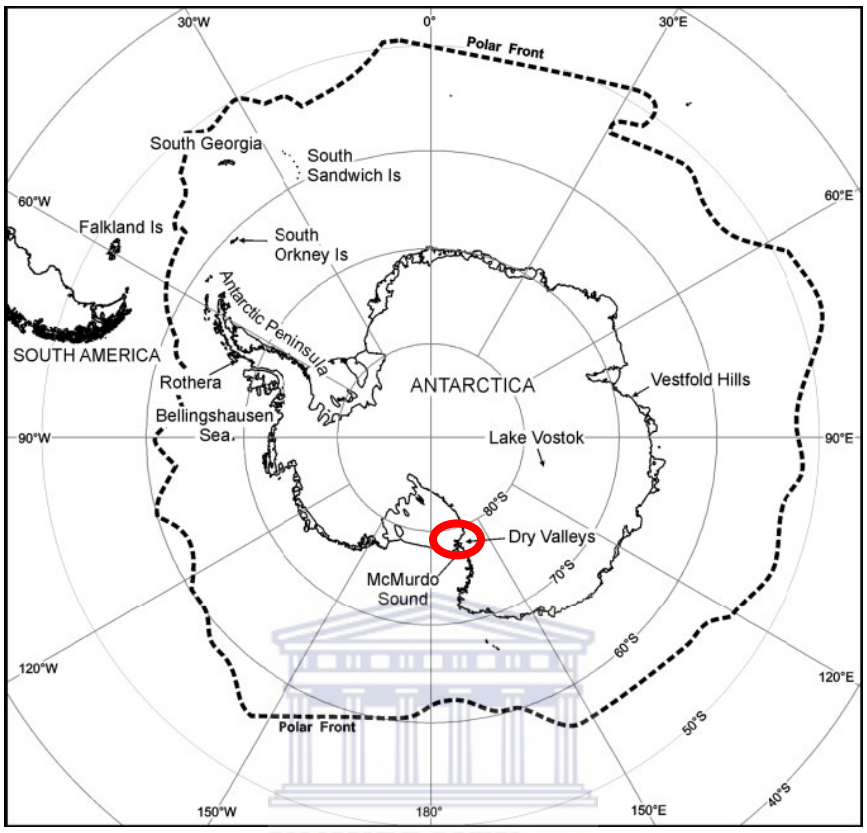




**Figure 1.1: Map of Antarctica continent and the Southern Ocean (Clark *et al.,* 2004).** The position of the McMurdo Dry Valleys is highlighted.

Antarctica (Fig. 1.1) is considered to be the largest, most pristine and extreme wilderness on our planet. It covers an area of nearly 4 million square kilometers and is the coldest, windiest and the highest continent on earth. Although much of the continent is ice covered, some coastal areas remains ice-free. The most significant ice-free region is the McMurdo Dry valley region on the eastern margins of the Ross Sea (Doran *et al*., 2002). Conditions in the Dry Valleys not only include low temperatures but also low water availability and nutrient deficiency (Friedman, 1993). These factors contribute to the limited colonization

of Antarctic biotopes. Antarctica contains 90% of the world's ice and approximately 70% of the world's freshwater.

Antarctic soil biotopes include the Ross Desert (McMurdo Dry Valleys) are known as the 'true desert' of Antarctica. Other major ice-free valleys of the Ross Desert include the Taylor, Wright, McKelvey, Balham, Victoria and the Barwick Valleys. Smaller valleys include the Miers (Fig. 1.2), Marshall, Garwood and Salmon Stream Valley (Simmons *et al.,* 1993; Vishniac, 1993). Soil biotopes are aerobic, and anaerobic microorganisms are considered to be very rare (Baker and Smith, 1972; Line 1988). The cold-active microorganisms that colonize these terrestrial habitats provide potential for biotechnological applications which include cold-active pectinolytic enzymes for use in the food industry and proteases, lipases and cellulases for use in detergents (Nakagawa *et al.,* 2004; Russell and Hamamoto, 1998).



### 1.2.1 Antarctic Dry Valleys

The McMurdo Dry Valleys, which are situated in the Eastern region of Antarctica, constitute the largest ice-free region on the continent (Doran *et al*., 2002). Conditions in these valleys include mean annual air temperatures of -20°C, with average temperatures of around 0°C during the summer but dropping to -55°C in the winter. During periods of direct sunlight the average ground surface temperature is around 15°C with the air temperature rarely rising above 0°C (Wynn-Williams, 1990). Due to the desiccating atmosphere, a low water content

of 0.5% to 2% wt can be expected in the upper soil layers. The desiccation of the soil surface is strongly increased by the katabatic winds that blow across the valleys during winter months. These winds reduce the humidity of the valleys to less than 10% (Horowitz *et al*., 1972). The average precipitation of the Dry Valleys is 15 g $cm^{-1}$ $year^{-1}$, in the form of snow. However very little moisture reaches the valley surface due to the high sublimation rate and very low atmospheric humidity.

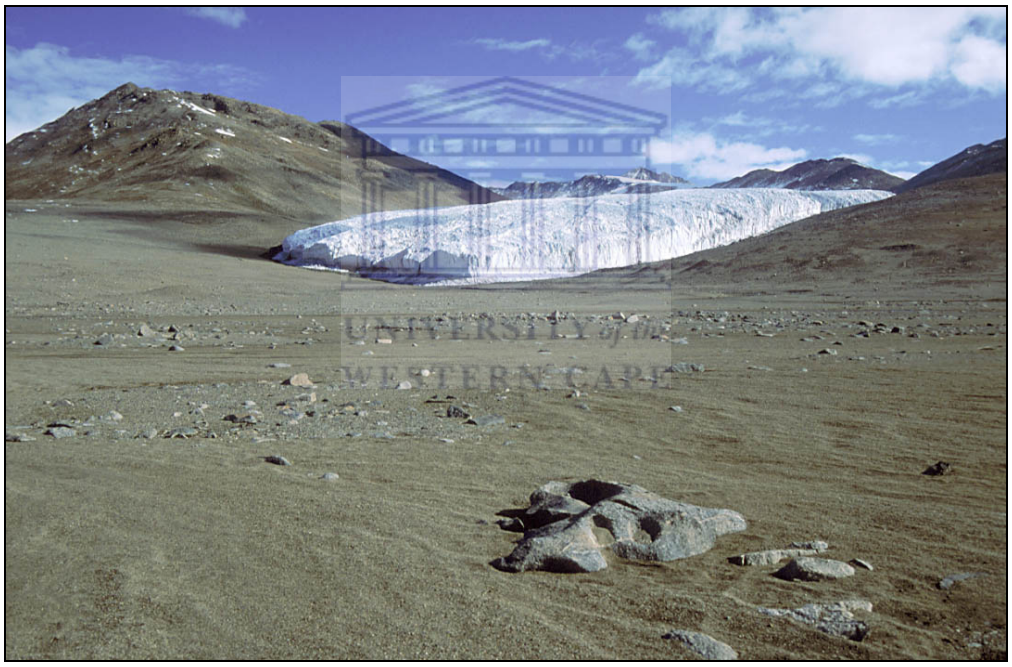

**Figure 1.2: The Miers Valley.**

The mineral soils of the Dry Valleys have high concentrations of mineral salts such as sodium, calcium, magnesium, chloride, sulfate and nitrate. The upward translocation from the substratum by capillary action and seaspray blown inland

by onshore winds are the main cause for the accumulation of these salts (Vishniac, 1993).

The Dry Valley mineral soils are very nutrient depleted because of the absence of plant life or any other major source of organic matter. Organic material would improve the biotic properties of Dry Valley mineral soils by facilitating water retention and reducing the high sublimation rate (Smith and Tearle, 1985). The extreme low water content in the McMurdo Dry Valley mineral soils was thought to give rise to low microbial numbers because water is known as a limiting factor in bacterial growth (Horowitz *et al*., 1972). McMurdo Dry Valley soils contains very low levels of organic carbon (≤0.05%), less than some mineral soils that contain between 5% and at most 15% organic carbon (Cameron, 1971). The sources of organic carbon in the Dry Valley are thought to be exogenous because not many chemoautotrophs have been documented. Although cyanobacteria and algae have also been isolated from Dry Valley soils, photosynthetic productivity is thought to be restricted to cryptoendolithic and hypolithic habitats, and in aquatic systems (streams, lakes, and flushes) (Friedmann, 1993). It is thus assumed that organic matter influx in the Dry Valleys is largely airborne-driven, from surrounding aquatic habitats. Kappen and Straka (1988) indicated that cyanobacterial mats from ice-covered lakes in Taylor Valley are dispersed aerially at an estimated rate of $2.93 \times 10^4$ kg.y$^{-1}$. The normal soluble nitrogen (N) concentrations measured in desert soils range from 0 to 1250 μg.g$^{-1}$, in the form of $NO_3^-$N, and from 0.3 to 40 μg.g$^{-1}$ in the form of $NH_4^+$ N (Cowan and Ah Tow, 2004). Antarctic mineral soils have one of the lowest

reported nitrate concentrations for terrestrial soils and acquire the nitrogen primarily from atmospheric precipitation (Cameron, 1971).

### 1.2.2 Microbial biomass in Antarctic desert soils

Quantitative data from culture-dependant studies on Dry Valley soils have indicated very low microbial numbers, ranging between $10^2$ to $10^4$ per gram of soil (Cameron, 1969; Cameron, 1971 & Cameron and Devaney, 1970). However, *in situ* ATP analysis data have shown that these soils contain three to four orders of magnitude higher levels of microbial biomass than indicated by quantitative data (Cowan *et al.,* 2002). This result supports the view that culture-dependant studies only access a minute fraction of the soil microbial flora present in the Dry Valley mineral soils. The highest numbers of culturable isolates in the McMurdo Dry Valleys have been found present in the permafrost layer and towards the surface (Cowan *et al.,* 2002; Tindall, 2004; Steven *et al.,* 2006).



### 1.2.3 Microbial diversity in Antarctic desert soil biotopes

Initial studies in the McMurdo Dry Valleys were based on classical culture-dependent methods or microscopy analysis (Cowan and Ah Tow, 2004). However, most investigations were done on aquatic systems, leading to a poor understanding of terrestrial Antarctic ecosystems. However, Cameron et al. (1972) observed that non-pigmented bacteria were mainly located below the soil surface whereas the soil surface was dominated by chromogenic bacteria. Plate isolations revealed both cosmopolitan and indigenous fungal, yeast, and

protozoan species from Dry Valley mineral soils. Most of these organisms were obtained from moist Dry Valley soils (Cameron *et al.,* 1972). In all previous studies it was observed that only a few microbial species in the Dry Valley soils were endemic and the majority showing a cosmopolitan distribution. Commonly isolated bacteria from McMurdo Dry Valley soils are represented in Table 1.1. Coryneforms were the most abundant bacteria isolated from Dry valley mineral soils of which 85% were represented by *Corynebacterium sepedonicum* (Cameron *et al.,* 1972; Cowan and Ah Tow, 2004).

**Table 1.1: Bacterial genera isolated from McMurdo Dry Valley mineral soils (summarized from Cameron *et al.,* 1972; Cowan and Ah Tow, 2004)**

| Microorganism | Genera |
|---|---|
| **Bacteria** | *Achromobacter* |
| | *Arthtrobacter* |
| | *Bacillus* |
| | *Corynebacterium* |
| | *Flavobacterium* |
| | *Micrococcus* |
| | *Planococcus* |
| | *Pseudomonas* |
| | *Streptomyces* |
| | *Nocardia* |



With recent advances in the field of molecular biology using DNA to investigate microbial diversity in various ecosystems it was found that 99% of microorganisms were not detected by culture-dependent studies (von

Wintzingerode *et al.,* 1997). In these investigations conserved phylogenetic markers that remain largely unchanged throughout the evolution of the microorganisms are used. The small subunit (ssu) rRNA gene is the molecule of choice in most phylogenetic surveys. This method proved to be useful in assessment of soil microbial communities identifying viable, nonviable, and dormant microorganisms of which most were found to be uncultured microorganisms. Most phylogenetic surveys of microbial species were done on temperate and tropical soils (Gentry *et al.,* 2006) and until recently; few studies have focused in the phylogenetic diversity of microorganisms in Antarctic Dry Valley mineral soils. Cowan and Ah Tow (2004) summarized the results from a phylogenetic study based on the 16S rRNA gene in the Miers Valley mineral soils and soils collected near the McMurdo station. These results are summarized in Table 1.2.



**Table 1.2: Bacterial genera identified based on 16S rRNA genes from Miers Dry Valley mineral soils and soils from McMurdo station (Cowan and Ah Tow, 2004)**

| Microorganism | Sample site + percentage sequence make-up | |
|---|---|---|
| | ***Miers Valley*** | ***McMurdo station*** |
| **Bacteria + Genera** | *Uncultured bacteria (50%)* | *Proteobacteria (37.5%)* |
| | *Actinobacteria (27%)* | *Bacteriodetes (25%)* |
| | *Bacteriodetes (11%)* | *Actinobacteria (12.5%)* |
| | *Acidobacteria (6%)* | *Firmicutes (12.5%)* |
| | *Verrucomicrobia (6%)* | *Plantctomycetes (12.5%)* |

In a recent phylogenetic survey in Dry Valley mineral soils by Smith et al. (2006) based on the 16S rRNA gene, eight phylogenetic groups were identified. Uncultured bacteria represented 21% of the phylotypes identified. The majority of these groups were Cyanobacteria, Actinobacteria and Acidobacteria. It was observed that the cyanobacterial phylotypes were unique to the desiccated high-altitude soils of the Dry Valleys which may suggest a soil-borne community. Similar observations were made by Aislabie et al. (2006) who identified a cyanobacterium-dominated cryptoendolithic community in Beacon sandstone. In comparison, Niederberger et al. (2007) identified cyanobacterial 16S rRNA signals only in samples with significant water availability.



### 1.3 Cyanobacteria

According to fossil evidence Cyanobacteria were in existence at least 3.5 billion years ago (Schopf, 1996). This fossil record includes both unicellular and filamentous forms of the group. Cyanobacteria are known to have unique phenotypic characteristics which include extracellular polysaccharides, production of UV-absorbing pigments, and phototactic motility. These characteristics are believed to be of great advantage to Cyanobacteria living on the primitive earth (Adams and Duggan, 1999). They show a very long evolutionary history in comparison to other microorganisms and play a significant role in the history of earth as primary producers (Tomitani *et al.,* 2006).

Cyanobacteria are a group of gram-negative prokaryotes (Stanier, 1988). They are morphologically diverse, a rare feature among prokaryotes, with some of

these organisms being unicellular and some multicellular (Rudi *et al.,* 2000). Their cell size varies from < 1 μM to > 100 μM in diameter. The thickness of their peptidoglycan layer together with its degree of cross-linking, and the presence of covalently linked polysaccharide are more similar to gram positive organisms. Their cell walls consist of two unit membranes, the cytoplasmic membrane and the outer membrane which is separated by the electron-opaque peptidoglycan layer (Stanier, 1988; Adams and Duggan, 1999; Rudi *et al.,* 2000). This cell wall also contains small diameter pores (5- 13 nm) which vary greatly from organism to organism.

Cyanobacteria are photoautotrophs that possess the ability to synthesize chlorophyll *a* (Whitton and Potts, 2000). Cyanobacteria contain photosystems I and II, the pathways of photosynthesis similar to the type found in the chloroplast of eukaryotic algae and higher plants. They also produce light harvesting proteins called phycobilin pigments designated phycocyanin (blue) and phycoerythrins (red).

Some members of the Cyanobacteria are also capable of anoxygenic photosynthesis, in which $H_2S$ is the electron donor rather than $H_2O$. The photosynthetic machinery is contained within the thylakoid membranes (Castenholz and Waterbury, 1989; Adams and Duggan 1999).

The nutrition of cyanobacteria is relatively simple because vitamins are not required, and either nitrate or ammonia is used as a nitrogen source. Some cyanobacterial species also have the ability to fix dinitrogen (diazotrophs) and

play a major role in the nitrogen economy of terrestrial habitats. Some cyanobacteria can assimilate simple organic compounds such as glucose and acetate in the presence of light through a process called photoassimilation (Madigan *et al.,* 2000).

The phylogenetic status of the Cyanobacteria was changed from eukaryotes (blue-green algae) to prokaryotes in the 1970s (Adams and Duggan, 1999). Cyanobacteria are grouped into five subsections based on morphological characteristics (Table 1.3). Members of subsection I, are known as the *Chroococcales* phylotypes. Organisms from this group are unicellular coccoid cell which may aggregate and divide by binary fission (Whitton and Potts, 2000). Members of subsection II, known as the *Pleurocapsales* are also unicellular coccoid but can also undergo multiple fission to produce smaller daughter cells in a capsule. These small daughter cells can easily disperse and are known as baeocytes (Tomitani *et al.,* 2006). Cyanobacteria of subsections III to V form filaments that differ morphologically. Filamentous Cyanobacteria from subsection III (*Oscillatoriales*) contains only vegetative cells that can divide only in a single plane. Filamentous organisms in subsections IV (*Nostocales*) and V (*Stigonematales*) are different from those in subsection III because they can produce vegetative cells that can differentiate into morphologically distinct heterocysts (cells that specialized in nitrogen fixation in aerobic environments). They can also form akinetes which are latent cells that can survive environmental stresses such as cold and desiccation. In addition to specialized cells,

filamentous organisms from subsection V have very complex branching patterns because they divide in more than one plane. This diversity and complexity of cyanobacterial groups are thus among the most highly developed in prokaryotes (Tomitani *et al.,* 2006; Whitton and Potts, 2000).

**Table 1.3: Cyanobacterial phylotypes and morphological characteristics (summarized from Tomitani *et al.,* 2006)**

| Subsection | Order | Characteristics |
|---|---|---|
| I | *Chroococcales* | Unicellular coccoid which divide by binary fission |
| II | *Pleurocapsales* | Unicellular coccoid which divide by multiple fission to produce smaller daughter cells in a capsule called baeocytes |
| III | *Oscillatoriales* | Filamentous (multicellular) vegetative cells that can divide only in a single plane |
| IV | *Nostocales* | Filamentous (multicellular) vegetative cells that form branches and heterocysts are present |
| V | *Stigonematales* | Filamentous (multicellular) vegetative cells that divide in more than one plane to form true branches with heterocysts and akinetes present |



The base composition of DNA from a variety of cyanobacterial species has been determined and shows that unicellular forms vary widely, from 35 to 71 % GC. However, the values for the filamentous forms vary much less, ranging from 38 to 46 % GC (Madigan *et al.,* 2000).

Cyanobacteria are widely distributed in nature from aquatic to terrestrial environments, which includes extreme habitats such as hot springs, deserts, and polar regions (Gallon *et al.,* 1996). Some members of the group can form symbiotic associations with different eukaryotic partners, including plants. Large numbers of Cyanobacteria are found in oceans where they are the primary producers. However, their distribution in soil environments is poorly understood. Members of this group were isolated from extreme environments ranging from high temperatures (>72°C) to high salinity (3M to 4M salt) (Reed *et al.,* 1984).

### 1.3.1 Cyanobacteria in Antarctic cold desert soils

The first description of cyanobacteria in Antarctica was by Griffith Taylor during the 1911 British expedition to the Southern Victoria land. Taylor observed the existence of “water plants” through the clear ice cover of the McMurdo Dry Valley lake (Taylor, 1916). Since Taylor’s early observation, most studies of Cyanobacteria in Antarctica have focused on cyanobacterial mats from lake, stream and pond systems (Brinkmann *et al.,* 2007). Wing and Priscu (1993) observed elevated levels of particulate organic carbon, organic nitrogen, ammonium, and dissolved organic carbon in lake ice and hypothesised that Cyanobacteria grow actively and not passively in the ice cover. Antarctic Cyanobacteria are well adapted to the extreme environment in terms of temperature, freeze-thaw survival, photoprotection and light acquisition for photosynthesis (Nadeau *et al.,* 1999). George et al. (2001) reported that the mucilage sheath produced by Antarctic Cyanobacteria played a role in

attachment to soil substrates, protection against desiccation, nutrient starvation and prevention against UV damage.

Despite the fact that Cyanobacteria dominate many Antarctic ecosystems, very few are considered as true psychrophiles (Tang *et al.,* 1997). They are best classified as psychrotolerant or psychrothrophic due to their high growth and photosynthetic optima (>20°C) and ability to metabolize near 0°C (Fritsen *et al.,* 1998). Interestingly, the same cyanobacterial species consisting mostly of *Phormidium*, *Oscillitoria* and *Nostoc* occur across habitats with different environmental extremes such as low temperatures, nutrient deficiencies and low water contents. However despite several studies on cyanobacteria in Antarctic marine systems, the terrestrial Antarctic Cyanobacteria have received little attention (Brinkmann *et al.,* 2007).



**1.4 Molecular methods used in microbial ecology**

The study of prokaryotic biodiversity in nature has been limited for many years due to the inability to access the true diversity of using culture dependant methods (Xu, 2006). Less than 0.1 % of the total microbial population can be isolated in pure culture using appropriate growth media and conditions (Nemergut *et al*., 2004). The problem arises because of the highly selective nature of the growth requirements of bacteria. In order to overcome this problem, a variety of media, culture conditions and other variables need to be used to obtain diverse microbial populations. However, this step becomes extremely laborious and time consuming (Xu, 2006).

Advances in molecular biological methods over the last decade have facilitated the analysis of uncultured microbial diversity (Fig. 1.3). These methods involve direct analysis of metagenomic DNA extracted from microbial communities. This analysis may involve techniques such as: Polymerase chain reaction (PCR), DNA cloning into suitable vectors, and community fingerprinting with denaturing gradient gel electrophoresis (DGGE) and/ or restriction fragment length polymorphism (RFLP) (Pace, 1986).

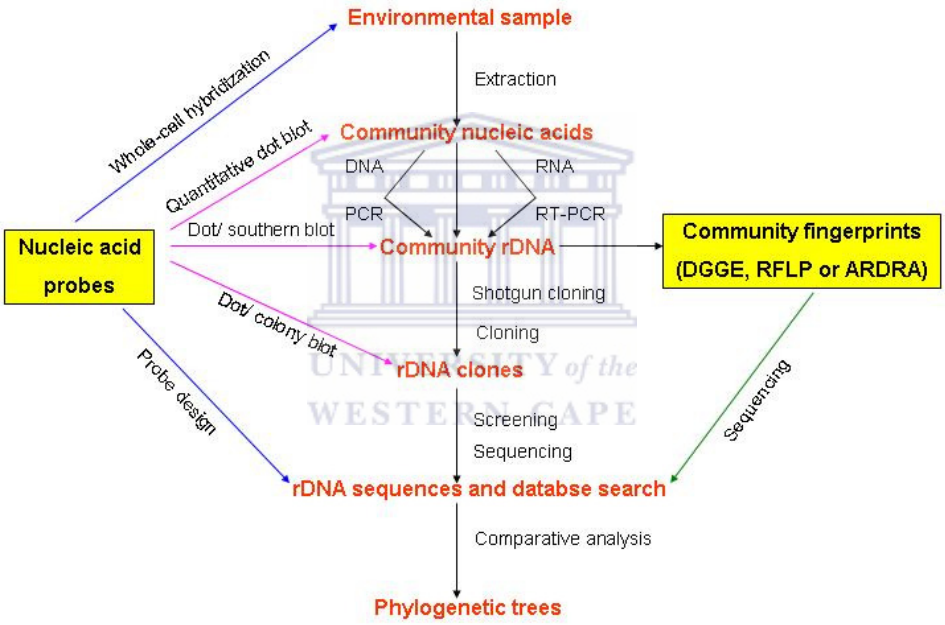


**Figure 1.3: Procedure for phylogenetic analysis of microbial populations in soils (modified from Pace, 1986).**

### 1.4.1 Sample collection and metagenomic DNA extraction

The collection of samples is often ignored when searching for the source of problems and pitfalls in molecular ecology (Liesack and Stackebrandt, 1992). However, this step is the start of metagenomic analysis and is crucial for all

subsequent steps. Collection effort and sample storage are important variables, and will normally be different from site to site. For example, small sample volumes such as soil samples can be stored on ice, frozen or processed immediately. This differs from marine environments which require extensive efforts in sampling, storage or processing on site (Liesack and Stackebrandt, 1992; von Wintzingerode *et al.,* 1997). However, special care must be taken during the transportation of samples from site to laboratory to avoid loss of nucleic acids due to freeze-thawing which causes lysis of cells (von Wintzingerode *et al.,* 1997). One way to overcome this problem may be to release and stabilize the nucleic acids immediately after the sample is collected (Muralidharan and Wemmer, 1994).



The extraction of metagenomic DNA from the environment is often a critical step when PCR- mediated approaches are being used for analysis (Xu, 2006). Only cells that lyse properly will form part of the final analysis of the diversity for that environment (von Wintzingerode *et al.,* 1997). Failure to lyse certain cell types will bias the apparent composition of community structure. Another problem that hinders metagenomic DNA extraction is the poor disruption of Gram-positive cells present in the sample. Over lysis of these cells may lead to highly fragmented DNA. The lysis efficiency may also be affected by the composition of the soil samples because microorganisms tend to adhere to some soil types (Xu, 2006). Different methods have surfaced to make DNA extraction from soil samples more effective. Two types of methods are employed for optimal DNA

extraction; cell extraction and direct lysis. Cell extraction involves the separation of whole cells from the sample prior to lysis and DNA extraction. However, this method has proved to be time consuming and includes laborious culture methods to obtain the whole cells before DNA extraction (Miller *et al.,* 1999). Direct lysis has proved to be much more effective because no cell separation is required before extraction (von Wintzingerode *et al.,* 1997; Xu, 2006). Direct lysis methods include bead-beating, where cells and soil particles are fragmented by vigorous agitation in the presence of glass beads (Miller *et al.,* 1999) and phenol chloroform extraction of DNA (Zhou *et al.,* 1996). Other methods involve enzymatic lysis using lysozyme and rapid freezing and thawing (Tsai and Olsen, 1991). However, a major disadvantage of these methods is the co-extraction of other organic compounds such as humic and fulvic acids. These contaminants inhibit further reactions, especially PCR. In such cases further purification protocols must be employed such as gel extraction, polyvinylpolypyrolidone (PVPP) or Sephacryl treatment (von Wintzingerode *et al.,* 1997).



### 1.4.2 Using the small-subunit ribosomal RNA (rRNA) gene as a phylogenetic marker

A number of biomolecules exist in nature which are potential phylogenetic markers. These molecules assist in the typing and grouping of different species according to similar genetic information (Urakawa *et al.,* 2003). Of these, it is widely accepted that the ribosomal RNA (rRNA) genes are uniquely suited for molecular phylogenetic analysis.

Several factors make rRNA genes, and particular the 16S, or small-subunit rRNA (ssurRNA) (Fig. 1.4), ideal for the analysis of evolutionary relationships (Nielsen *et al.,* 2003): (1) All organisms posses rRNAs, making a universal phylogeny possible. (2) The lateral transfer of rRNAs between different organisms is extremely rare. (3) The longer rRNAs (16S, 18S and 23S) contain regions of highly conserved, moderately variable and highly variable sequences (Gentry *et al.,* 2006; Letowski *et al.,* 2004). The highly variable regions (hypervariable regions) are an indication of sequence divergence over evolutionary time (Neefs *et al.,* 1990). These variable regions are often flanked by highly conserved regions which act as alignment guides to ensure that only homologous sequences are compared between organisms (Gentry *et al.,* 2006). The comparison is normally made by using polymerase chain reaction (PCR) and the primers designed from conserved regions to amplify the hypervariable regions. This method has lifted the burden of cultivation of microorganisms before characterization (von Wintzingerode *et al.,* 1997) and emphasizes the huge impact that rRNA based methodologies have on the field of microbial ecology, especially in phylogenetic studies (Gentry *et al.,* 2006).

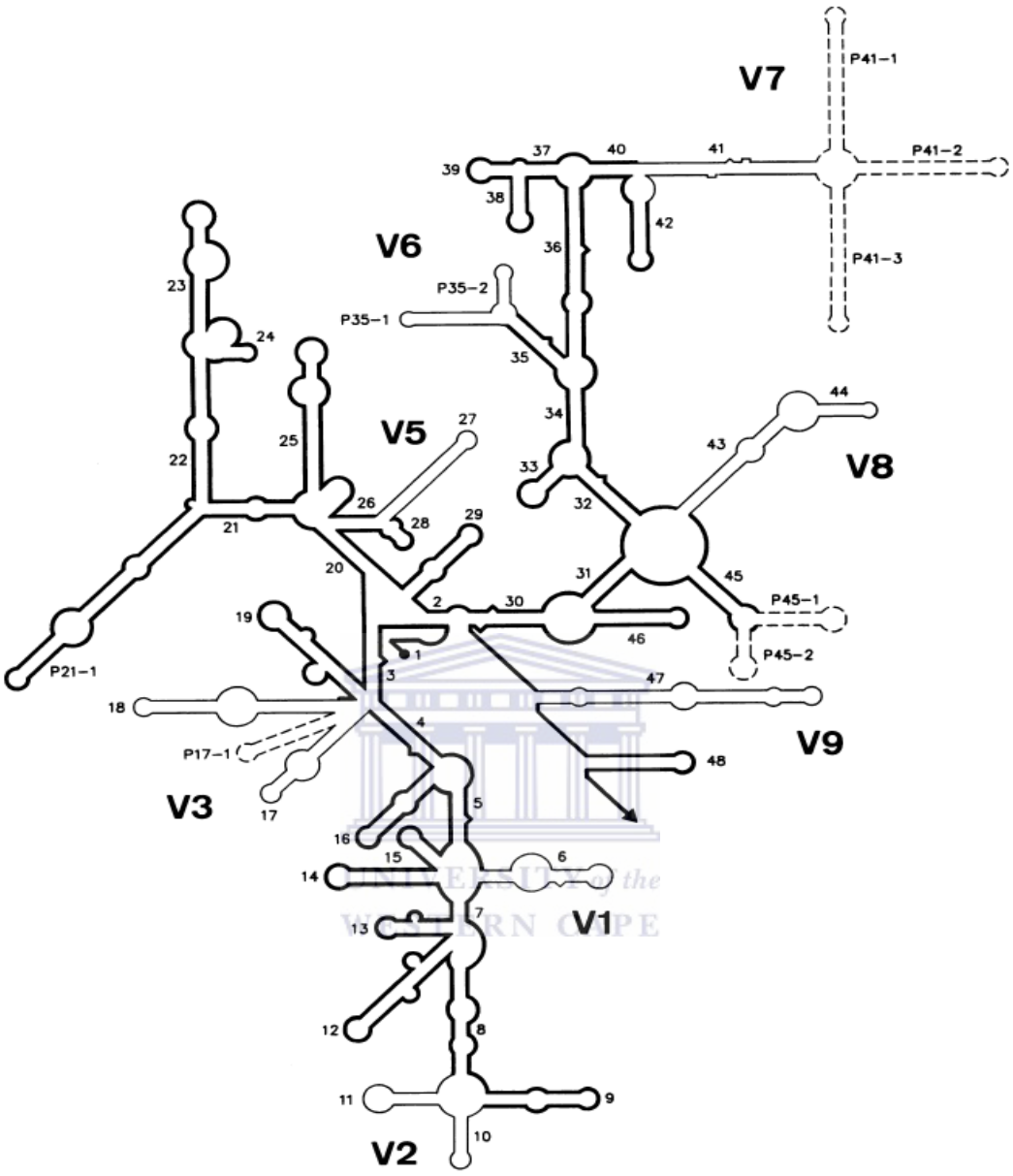


**Figure 1.4: Secondary structure for prokaryotic ssrRNAs. The 5'- terminus is symbolized by a filled circle and the 3'- terminus by an arrowhead. Helices are numbered in the order of occurrence from 5'- to 3'- terminus. Helices bearing a single number are common to the prokaryotic and eukaryotic models. A composite number preceded by P points to a prokaryotic-specific helix. Relatively conserved regions are drawn in bold lines, areas of sequence-and length variability in thin lines. Eight variable areas, numbered V1 to V9, are distinguished, V4 being absent in prokaryotic ssrRNA. Helices drawn in broken lines are present in small number of known structures only. Archaebacterial sequences follow the prokaryotic pattern except for helix 35, which is unbranched as in eukaryotic (Neefs *et al.,* 1990).**

### 1.4.3 DNA fingerprinting techniques

Different DNA fingerprinting techniques are used to differentiate between sequences prior to the sequencing of genes (von Wintzingerode *et al.,* 1997). These methods typically distinguish between sequences based on the absence or presence of specific restriction sites. These methods include using restriction enzymes for restriction fragment length polymorphism (RFLP), amplified ribosomal DNA restriction analysis (ARDRA) and terminal restriction fragment length polymorphism (T-RFLP). ARDRA is used to differentiate between 16S rDNA sequences after PCR amplification (Vaneechoutte *et al.,* 1993). The resulting products obtained after restriction digestion with the relevant enzyme are normally sized fractionated by gel electrophoresis (Bermingham and Luettich, 2003). The resulting restriction fragments represent the intraspecific differences of the conserved rDNA (Bermingham and Luettich, 2003).



Sequence-based DNA fingerprinting techniques include temperature gradient gel electrophoresis (TGGE), denaturing gradient gel electrophoresis (DGGE) and single stranded conformation polymorphism (SSCP) analysis. All of these methods provide a DNA fingerprint of sequences with high resolution for subsequent analysis (Oto *et al.,* 2006). In each of the methods PCR-amplified products are loaded onto polyacrylamide (PAA) gels for capillary electrophoresis. The products migrate to different locations on the gels based on sequence differences (Hayes *et al.,* 1999). In TGGE, the PAA gel is subjected to a temperature gradient and the gene products migrate based on the different melting properties of the sequences. This method is similar to DGGE, where a

gradient with increasing concentration of denaturant (urea or formamide) is used (Muyzer *et al.,* 1993). As the double stranded DNA (dsDNA) is electrophoresed through the denaturing gradient, it melts and changes conformation. This change will affect the mobility of the dsDNA based on the principle that A-T rich fragments have lower melting temperature than G-C rich fragments. The use of a G-C clamp (a 30-40 base pair sequence of G-C nucleotides) on one of the primers is critical for DGGE, because it stabilizes the melting fragments and prevents complete strand dissociation (Hayes *et al.,* 1999; Muyzer *et al.,* 1993).

SSCP differs from DGGE because neither a denaturing gradient nor a G-C clamp is used in the analysis (Orita *et al.,* 1989). Another difference is the conformation change of the DNA through the PAA gel. In SSCP the DNA strands gets separate by high temperature and the single strands migrate to different positions in the gel based on conformation (Hong *et al.,* 2007).



All of these methods can detect up to 96% of all mutations or single-point mutations present in the genes targeted.

### 1.4.3.1 Applications of DGGE in molecular ecology

The DGGE method is extremely widely used in molecular ecology usually to asses the structure of microbial communities in environmental samples (Ercolini, 2004). The method is also used to determine community dynamics in response to environmental changes. The most recent application of this method is in the study of soil microbial communities to evaluate structure and evolution of these communities (Avrahami *et al.,* 2003). Other applications include marine

environments (Bano and Hollibaugh, 2002); rivers (Sekiguchi *et al.,* 2002), and lake waters (Crump *et al.,* 2003). DGGE has also been used to identify and type specific microbial species from the different environments by using specific primers to target selected microbial species. Microbial studies are not the only application of the DGGE technique; it can also be used to identify genes from the environment (Ercolini, 2004). In these types of studies the DGGE-primers are designed to target specific genes and the genes are targeted using PCR amplification following by DGGE (Jackson *et al.,* 2000).

## 1.5 Aims of this study

The Antarctic mineral soils have been broadly studied using culture-dependant approaches (Friedmann, 1993). However, recent advances in molecular phylogenetics has allowed for the analysis of microbial diversity in all biotopes including other Antarctic niches such as cryptoendolithic (de la Torre *et al*., 2003), cryoconite holes (Christner *et al*., 2003) and lake ice (Priscu *et al*., 1998) and lake (Brambilla *et al*., 2001) communities. Antarctica has been used as a model in the quest for extraterrestrial life (Friedmann, 1993). Conditions such as extreme desiccated soils with very low water contents and high UV incidence have driven the suggestion that Antarctic ice-free regions are similar to that of Mars. These ice-free deserts harbour many microorganisms specialized to survive in the harsh conditions such as Cyanobacteria, Firmicutes, Proteobacteria, Actinobacteria and Bacteriodetes (Cowan and Ah Tow, 2004). Cyanobacteria have been proposed as the dominant bacteria phylum in Antarctic biotopes and have been extensively studied in Antarctic aquatic environments such as lakes and ponds (Tang *et al.,* 1997). However, the terrestrial Antarctic cyanobacteria have received little attention (Brinkmann *et al.,* 2007). Here we attempt to better understand the cyanobacterial community structure in Antarctic Dry Valley mineral soils by using PCR-based strategies to obtain cyanobacterial 16S rDNA sequence information from extracted metagenomes to investigate their distribution in the Antarctic Miers Valley. Specific aims were**:**

- To isolate high molecular weight DNA of considerable purity from eleven sample sites collected along the northern transect in the Miers Valley, Eastern Antarctica.

- PCR amplification of cyanobacterial16S rRNA genes.

- Constructing a cyanobacterial community profile using DGGE analysis.

    - To use this DGGE profile to study the diversity of cyanobacterial phylotypes in Antarctic Dry Valley mineral soils.

    - Study the effect of factors such as spatial separation, altitude, soil texture and relative soil humidity have on cyanobacterial community structure.

- To construct cyanobacterial 16S rDNA libraries from four selected sites.

    - To use amplified ribosomal DNA restriction analysis (ARDRA) to discriminate between 16S rDNA genes from the different libraries for sequencing and gene identification.

    - Sequence data will provide information of which cyanobacterial phylotypes are specific to moist or extremely dry soils in the Dry Valleys.

# CHAPTER TWO

## Materials and Methods

### 2.1 Reagents used

All bacterial strains and plasmids used in this study are listed in Table 2.1. Unless otherwise stated, all chemicals used were obtained from Sigma Chemical Co or Merck Laboratory Supplies and of analytical grade. Other materials used in this study were obtained from Merck (Biolab Diagnostics), Oxiod (Oxoid Ltd) or Kimix Laboratory Supplies.

**Table 2.1: Bacterial strains and plasmids used in this study**

| Strains & Plasmids | Characteristics | Reference |
|---|---|---|
| **Strains:** | | |
| ***E. coli*** | | |
| DH10B | F⁻ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*Δ139 Δ(*ara*, *leu*)7697 *gal*U *gal*K λ- *rps*L (Str$^R$) *nup*G | Invitrogen |
| | | |
| **Plasmids:** | | |
| pTZ57R/T | *Linearised with 5'-ddTs added, lacZ' Ap$^R$* | Fermentas |

### 2.2 Antibiotics

Addition of antibiotics to either liquid broth or growth media was performed aseptically. Broth or media were sterilized by autoclaving at 1.0 kg cm$^{-2}$, 121°C

for 20 min using an Austester-G autoclave and cooled to approximately 45°C before supplementation with the appropriate filter sterilized antibiotic. The final concentration of antibiotic in both media and agar was: ampicillin, 100 µg $ml^{-1}$.

## 2.3 Culture Media

### 2.3.1 Luria-Bertani (LB) broth

LB broth consisted of 1% (w/v) tryptone, 0.5% (w/v) yeast extract and 1% (w/v) NaCl. The broth was prepared by mixing together all constituents with distilled water and pH adjustment to 7.0 prior to autoclaving.



### 2.3.2 LB agar

LB agar consisted of LB broth prepared as in section 2.3.1 with the addition of 1.3% bacteriological agar prior to autoclaving.

### 2.3.3 SOB medium without magnesium

SOB medium without magnesium consisted of 2% (w/v) tryptone, 0.5% (w/v) yeast extract and 0.05% (w/v) NaCl and 2.5 mM KCl. All constituents were mixed together with distilled water and the pH was adjusted to 7.0 prior to autoclaving.

### 2.3.4 Terrific Broth (TB)

TB was prepared according to Sambrook and Russell (2001) and consisted of 1.2% (w/v) tryptone, 2.4% (w/v) yeast extract, 0.4% (v/v) glycerol, 17 mM $KH_2PO_4$ and 72 mM $K_2HPO_4$ in distilled water. Tryptone, yeast extract and

glycerol were prepared separately in 900 ml distilled water before autoclaving. All solutions were mixed at room temperature.

### 2.4 Enzymes

The following enzymes were used in this study: Restriction enzymes, T4 DNA ligase (Fermentas) and thermostable *Taq* DNA polymerase (in-house prepared according to Desai and Phaffle, [1995]).

### 2.5 Sample collection and storage

Antarctic soil samples were collected by D. A. Cowan along an altitudinal transect on the northern slope of the Miers Valley, Eastern Antarctica, during December 2005 and January 2006. Samples were collected from near Lake Miers at 171 m a.s.l, to the top of the slope 634 m a.s.l at 50 m increments. At each site a 1 m square was set up, and samples were collected at each of the four corners. Approximately 50% of each sub-sample was pooled, and a fifth homogeneous sample taken. Collection was carried out in sterile 50 ml Falcon tubes by removal of a 1 cm surface layer of mineral soil. All samples (approx. 400g each) were mixed thoroughly and resampled before storage at below 0°C for transport to the laboratory. Samples were stored at -80°C until required for further analysis.

## 2.6 Metagenomic DNA isolation

### 2.6.1 Modified Miller method

Metagenomic DNA extraction was done on the basis of the Miller protocol (Miller *et al.,* 1999). Between 0.5 and 1 g of soil was added to sterile 2 ml screw cap tubes containing 1 g of sterile quartz sand, followed by thoroughly mixing and the addition of 300 µl 0.1 M sodium phosphate, pH 8, 300 µl lysis buffer (0.5 M Tris-HCl, pH 8, 10% SDS [v/v], 100 mM NaCl) and 300 µl chloroform/ isoamyl alcohol (24:1). The mixture was subjected to freeze-thawing by incubation at 80°C in an AccuBlock™ Digital Dry bath (Labnet International, Inc) for 20 min followed by cooling on ice for 15 min. The sample tubes were vortexed for 2 min at maximum speed, followed by centrifugation for 10 min at 15 000 x *g*. The Supernatants were transferred to sterile 1.5 ml Eppendorf and ammonium acetate was added to a final concentration of 2.5 M. The tubes were inverted several times until white flocculates appeared, then centrifuged for 10 min at 15 000 x *g*. The supernatants were collected and transferred to 1.5 ml Eppendorf tubes followed by addition of 0.6 volumes of Isopropanol. After inversion several times, the tubes were incubated at room temperature for 30 min. DNA was pelleted by centrifugation at 15 000 x *g* for 30 min at room temperature, washed with 70% ethanol and air dried in a sterile hood for 5 min. TE buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA [pH8.0]) was used to dissolve the DNA pellet followed by storage at -20°C until further analysis.

### 2.6.2 DNA quantification

All DNA concentrations were measured using a Qubit™ Flourometer (Invitrogen) according to the manufactures instructions. DNA purity was measured using the Nano-drop ND-1000 according to the manufactures instructions. In both instruments, 2 µl of DNA was used for analysis.

## 2.7 Agarose gel electrophoresis

DNA fragments were separated and visualized using agarose gel electrophoresis (Sambrook and Russell, 2001). The appropriate percentage (v/w) TAE agarose gels were cast and electrophoresed at 100 V in 0.5X TAE buffer (20 mM Tris, 10 mM sodium acetate and 0.5 mM EDTA [pH 8.0]). For DNA recovery from agarose gels, 0.7% - 1% (w/v) TAE agarose gels were cast and electrophoresed at 60 V in 0.5X TAE buffer. Visualization of DNA was done using an AlphaImager 2000 digital imaging system (Alpha Innotech, San Leandro, CA) after supplementing the agarose gels with 0.5 µg $ml^{-1}$ ethidium bromide. DNA size determination was done by comparing migrated DNA fragments on the agarose gels to standard molecular markers (self-made DNA ladder or 100 bp ladder [Roche]).



## 2.8 GFX™ DNA purification

Purification of DNA excised from agarose gels were performed using the GFX™ PCR and Gel Band Purification Kit (Amersham Biosciences) according to manufacturer's specification.

## 2.9 Restriction endonuclease digestion

All restriction enzyme digestions were performed in either sterile Eppendorf tubes or in 96-well microtiter plates in small reactions volumes (10 – 20 µl). Reactions contained the appropriate volume of 10X specific enzyme buffer supplied by the manufacturer, and 2.5 – 5U of enzyme per µg of DNA. Reactions were incubated for 2 h in a waterbath or in Gene Amp® PCR machines at 37°C. The digestion products were analysed by agarose gel electrophoresis in TAE agarose gels as described in Section 2.7.

## 2.10 PCR amplifications

All primers and their optimal annealing temperatures are shown in Table 2.2. Target DNA was amplified using 0.2 ml thin walled tubes in Gene Amp®, Eppendorf Master cycler gradient or Thermo Hybaid PCR Sprint machines. PCR reactions (25 – 50 µl) contained 10 ng of metagenomic DNA, 1X PCR buffer (from 10X buffer consisting of 200 mM Tris [pH 8.8], 100 mM KCl, 100 mM $[NH_4]_2SO_4$, 20 mM $MgSO_4$, and 1% [w/v] Triton-X-100), 0.5 µM of both forward and reverse primers, 200 µM dNTP mixture (dATP, dCTP, dGTP and dTTP), and 1.5 U *Taq* DNA polymerase. Reactions were made up to the appropriate final volume using $ddH_2O$. An aliquot of the PCR reaction mixture was analysed by gel electrophoresis as described in Section 2.7.

Unless otherwise stated thermocycling conditions were as follows: Initial denaturation at 94°C for 4 min, followed by 30 cyles of denaturation at 94°C for

30 s, annealing at X°C for 30 s, and extension at 72°C for 1 min (X denotes the primer-specific annealing temperature as indicated in Table 2.2). A final elongation step of single cycle was performed at 72°C for 10 min.

### 2.10.1 Colony PCR

DNA templates for colony PCR were prepared by resuspending bacterial colonies in 5 µl dd$H_2O$. The cells were lysed by boiling at 100°C for 5 min followed by centrifugation at 16 000 x *g* for 5 min at 4°C. The supernatants were collected and 1 µl was used as DNA template in a 25 µl PCR reaction (Section 2.10). The primer pair M13fw and M13rev (Table 2.2) was used at annealing temperature of 53°C with cycling conditions as in Section 2.10.



### 2.10.2 Nested PCR

Where required, nested PCR was performed using first round PCR amplification products (as in Section 2.10 or 2.10.1) as DNA templates. The second round amplification was done using 1 µl of product from the first round amplification as the template in a 50 µl PCR reaction.

**Table 2.2: Primers used for PCR amplifications in this study**

| Primers | Sequence (5'-3') | Annealing temperature (X) | Reference |
|---|---|---|---|
| E9F | GAGTTTGATCCTGGCTCAG | 50°C | Hansen et al. (1998) |
| U1510R | GGTTACCTTGTTACGACTT | 50°C | Reysenbach et al. (1995) |
| M13fw | CCCAGTCACGACGTTGTAAAACG | 53°C | pTZ57R/T plasmid specific |

**Table 2.2: Continued**

| Primers | Sequence (5'-3') | Annealing temperature (X) | Reference |
|---|---|---|---|
| M13rev | AGCGGATAACAATTTCACACAGG | 53°C | pTZ57R/T plasmid specific |
| CYA359 (T)F | GGGGAATTTTCCGCAATGGG | 55°C | Boutte et al. (2005) |
| CYA359 (T)FGC | *GGGGAATTTTCCGCAATGGG | 64°C | Boutte et al. (2005) |
| CYA781 aR | GACTACAGGGGTATCTAATCCCTTT | 64°C | Nübel et al. (1997) |
| CYA781 bR | GACTACTGGGGTATCTAATCCCATT | 64°C | Nübel et al. (1997) |

*additional GC-clamp: CGCCCGCCGCGCCCCGCGCCCGTCCCGCCGCCCCCGCCCGGG

## 2.11 Construction of cyanobacterial 16S rDNA libraries

### 2.11.1 Preparation of *E. coli* electrocompetent cells

Electrocompetent *E. coli* DH10B cells (Table 2.1) were prepared as described by Sambrook and Russell (2001), with slight modification. All glassware was sterilised by acid washing with 30% $H_2SO_4$, followed by autoclaving prior to use. A single colony of DH10B was inoculated into 10 ml of LB-broth and incubated overnight at 37°C with shaking (250 rpm). 3 ml of the culture was transferred to two aliquots of 500 ml SOB medium without magnesium (Section 2.3.3) and incubated at 37°C with shaking until an optical density at 600 nm ($OD_{600nm}$) of 0.6 was reached. The cells were rapidly cooled on ice-water for 10 min and collected in polypropylene tubes by centrifugation in a Beckman J2-21 centrifuge at 4 000 x *g* for 25 min at 4°C. The supernatant was removed and the cells resuspended in 200 ml ice-cold $dH_2O$. Cells were centrifuged as above and the supernatant removed followed by resuspension in 100 ml ice-cold $dH_2O$. The cells were subjected to another centrifugation step (as above) and supernatants decanted followed by resuspension in 20 ml ice-cold 10% glycerol. Two 20 ml cell

suspensions were mixed in one chilled conical tube and harvested by centrifugation at 4 000 x *g* for 10 min. The supernatant was removed and the cell pellet resuspended in 1 ml ice-cold 10% glycerol to a final cell density at $OD_{600nm}$ of 250. The cells were aliquoted into 50 µl volumes, snap frozen in liquid $N_2$ and stored at -80°C until required.

### 2.11.2 PCR amplification of partial cyanobacterial 16S rDNA

PCR amplification targeting partial cyanobacterial 16S rDNA templates was performed as described in Section 2.10. PCR reactions of 50 µl contained 10 ng metagenomic DNA, as in Section 2.10. The primer pair of CYA359(T)F and U1510R (Table 2.2) were optimized and an optimal annealing temperature of 53°C was identified. This annealing temperature was used with cycling conditions as described in Section 2.10. PCR products were visualized by gel electrophoresis as described in Section 2.7. PCR products were excised from the agarose gel using the steps described in Section 2.7 and purified using the GFX™ PCR and Gel Band Purification Kits (Section 2.8) according to manufacturer's instructions. DNA quantification was done using the Qubit™ flourometer (Invitrogen) (Section 2.6) according to manufacturer's instructions.



### 2.11.3 DNA ligations

Ligation reactions were performed using the InsT/Aclone Kit (Fermentas) according to the manufacturer's instructions. Assays were prepared in 15 µl reactions consisting of the linearised plasmid (pTZ57R/T) (Table 2.1), insert DNA

(molar ratio, 1: 3), 1X ligation buffer, 1U T4 DNA ligase and $dH_2O$. Mixtures were briefly vortexed, centrifuged at 15 000 x *g* for 5 s and incubated overnight at 16°C.

### 2.11.4 Transformation of *E. coli* cells by electroporation

The *E. coli* DH10B electrocompetent cells (Section 2.11.1) were transformed by the addition of 1 µl ligated DNA (Section 2.11.3) to 50 µl thawed cells and gently mixed. The mixture was cooled on ice for 1 min, then transferred into a pre-cooled 0.1 cm sterile electroporation cuvette (Bio-Rad Laboratories, Hercules, CA, USA). Electroporation was performed at the following conditions: 1.8 kV, 25 µF, and 200 Ω. After electroporation, 950 µl TB broth, pre-warmed to 37°C, was added to the cuvette and properly mixed. The cells were transferred to a 2 ml Eppendorf tube and incubated at 37°C for 90 min with agitation. Cells were plated in aliquots of 5 to 100 µl onto LB-agar plates supplemented with 100 µg $ml^{-1}$ ampicillin, 20 µg.$ml^{-1}$ IPTG (iso-propyl-ß-thiogalactopyranoside), and 30 µg $ml^{-1}$ X-Gal (5-bromo-4-chloro-3-indolyl-ß-D-galactopyranoside) for blue/ white colour selection. The agar plates were incubated overnight at 37°C.



### 2.11.5 Screening of recombinant clones

Initial screening of recombinant clones for correct inserts following transformation (Section 2.11.4) was done using colony PCR as described in Section 2.10.1. Aliquots of the PCR reaction mixtures were analysed using gel electrophoresis on 1.5% TAE agarose gels as described in Section 2.7.

#### 2.11.5.1 Screening positive clones for cyanobacterial 16S rDNA

All positive clones screened with colony PCR were subjected to cyanobacterial 16S rDNA screening using a nested PCR approach (Section 2.10.2). The first round PCR amplicons from colony PCR (Section 2.11.5) were diluted 1:10 and used as DNA templates (1 µl) in a second round amplification of 25 µl per PCR reaction. The primer pair CYA359(T)F and an equimolar mix (ratio, 1:1) of CYA781aR and CYA781bR (Table 2.2) was used at an annealing temperature of 64°C with cycling conditions as described in Section 2.10. Aliquots of the PCR reaction mixture were analysed using gel electrophoresis on 1.5% TAE agarose gels as described in Section 2.7.

### 2.11.6 ARDRA analysis of 16S rDNA amplicons

PCR amplicons were prepared using a nested PCR approach (Section 2.10.2). First round PCR amplicons from colony PCR (Section 2.11.5) were diluted 1:10 and used as DNA templates (1 µl) in a second round amplification of 50 µl per PCR reaction according to Section 2.11.2. Aliquots of the PCR reaction mixture were analysed using gel electrophoresis on 1% TAE agarose gels as described in Section 2.7. Amplicons of the correct size were subjected to restriction endonuclease digestion (Section 2.10) using 8 µl of the PCR reaction, 2.5 U of the *Hae*III restriction enzyme and 5 U of *Rsa*I separately. Restriction digestion products (15 µl) were viewed on 3% agarose gels following gel electrophoresis (Section 2.7).

## 2.12 Nucleic acid sequencing

### 2.12.1 Plasmid extractions

Plasmid extractions performed for subsequent nucleotide sequence analysis were done using the Invisorb® Spin Plasmid Mini *Two* Kit (Invitek) according to the manufacture's instructions.

### 2.12.2 Sequencing

Sequencing of cloned insert DNA was performed using the MegaBace 500 Automated Capillary DNA Sequencing System (Amersham Biosciences) at the University of Cape Town sequencing facility (South Africa). The vector primers M13fw and M13rev (Table 2.2) were used depending on the insert direction for sequencing of the 16S rDNA library inserts. Closest matches of the partially sequenced 16S rDNA genes within the GenBank database were identified by BLAST analysis using the National Center for Bioinformatics Information (NCBI) server (http://www.ncbi.nlm.nih.gov/blast/). The Ribosomal Database Project (RDP) Classifier (http://rdp.cme.msu.edu/classifier/classifier.jsp) was used to assign unclassified sequences to taxonomic groups.



## 2.13 Denaturing gradient gel electrophoresis (DGGE)

### 2.13.1 Amplification of cyanobacterial 16S rDNA fragments for DGGE

Cyanobacterial amplicons for DGGE analysis were prepared according to the protocol of Nübel et al. (1997) with slight modification. A nested PCR strategy was used according to Section 2.10.2. The PCR reaction mixtures were prepared

as in Section 2.10. The first round reaction was done as in Section 2.11.2. First round PCR amplification products were used as DNA templates (1 µl) in a second round amplification of 50 µl per PCR reaction. The primer pair CYA359(T)F GC and equimolar mix (ratio, 1:1) of CYA781aR and CYA781bR (Table 2.2) were used at an annealing temperature of 64°C with cycling conditions as follows: Initial denaturation at 94°C for 4 min, followed by 35 cyles of denaturation at 94°C for 45 s, annealing at 64°C for 45 s, and extension at 72°C for 1 min. A final elongation step was performed at 72°C for 30 min. Aliquots of the PCR reaction mixture were analysed by gel electrophoresis on 1.5% TAE agarose gels as described in Section 2.7.

**2.13.2 DGGE**

DGGE was performed based on the protocol of Muyzer et al. (1993) using the Dcode™ universal mutation detection system (Bio-Rad) according to the manufacture's instructions. High- and low denaturing solutions (20 ml) were prepared before gel casting. High denaturing solutions were prepared according to Table 2.3. A 1mm thick 9% polyacrylamide gel (Acrylamide/bis 37.5:1) (16 x 16 cm) in 1x TAE buffer and the polymerization was initiated by adding 16 µl of TEMED and 160 µl 10% APS to both high- and low denaturing solutions. The gel prepared with denaturing gradient of 20-60% was poured between the assembled plates using a gradient former and allowed to polymerise for 90 min.

**Table 2.3: Preparation of the solutions for a 9% polyacrylamide DGGE gel with a denaturing gradient of 20-60%**

| UF denaturant | 20% | 60% |
|---|---|---|
| 40% Acrylamide | 4.5 ml | 4.5 ml |
| 50X TAE | 400 µl | 400 µl |
| UREA | 1.68 g | 5.04 g |
| Formamide | 1.6 ml | 4.8 ml |

*UF= Urea and Formamide

Before gel electrophoresis, 18 µl of PCR products of the second round (Section 2.13.2) were loaded onto the solidified DGGE gel. The gels were electrophoresed in 1X TAE buffer at a constant voltage of 100 V for 16 h at 60°C. After electrophoresis, the gels were stained for 15 min in 0.5 µg $ml^{-1}$ ethidium bromide solution and destained in 1X TAE. The DGGE gels were visualized using the AlphaImager 2000 digital imaging system (Alpha Innotech, San Leandro, CA).



### 2.13.2.1 DGGE cluster analysis

Cluster analysis was performed to investigate the relationships between the DGGE profiles. DGGE bands were manually scored as absent or present for the generation of a data set for analysis with the AlphaIEaseFC™ (version 4.0) software package (Alpha Innotech, San Leandro, CA). Using the pair-wise comparisons between profiles, a distance matrix was generated for using in Neighbour Joining analysis to determine how the profiles clustered.

# CHAPTER THREE

## Results

### 3.1 Introduction

Soil microbial communities are the most complex and highly diverse of all habitable environments on Earth (Kuske *et al.,* 1997). Reassociation kinetics estimates place bacterial numbers at approximately $10^9$ cells per g of soil and over $10^4$ different species (Dunbar *et al.,* 2002). However, the desiccated Antarctic Dry Valley mineral soils have relatively low bacterial numbers, which is a characteristic of extreme environments (Cameron *et al.,* 1972). ATP, lipid and DNA quantification analysis have shown that Antarctic soils may contain between $10^6$ and $10^8$ bacterial cells per g (Cowan *et al.,* 2002).



Until recently, Antarctic soil microbiology has been exclusively investigated using culture-dependent approaches and microscopy. These studies have shown that most Antarctic microbes belong to a specialized and restricted number of cosmopolitan taxa consisting largely of aerobic microbes (Smith *et al.,* 2006). Cyanobacteria, an ancient group of photosynthetic microorganisms are considered the most important and dominant photoautotrophs in terrestrial and aquatic Antarctic ecosystems (Taton *et al.,* 2003). The benthic proliferations of cyanobacteria in Antarctic lakes and ponds have attracted the most interest leading to research into their molecular and morphological taxonomy (Vincent *et al.,* 1993). Other studies include cyanobacterial community analysis, biogeographical distribution and physiology analysis in Antarctic aquatic

ecosystems (de los Rios *et al.,* 2004; Taton *et al.,* 2003). In contrast, the soil cyanobacteria communities of Antarctica have been poorly studied (Cavacini, 2001). Studies done on cyanobacteria in Antarctic soils have relied primarily on microscopic identification and culturing methods. This indicates the importance of research characterising these soils and investigating the parameters which influences cyanobacterial community structure (Taton *et al.,* 2003).

Recent advances in molecular biology have provided phylogenetic techniques to study complete bacterial communities in almost every environment. The construction of 16S rDNA clone libraries has provided detailed phylogenetic information of Antarctic cyanobacterial biodiversity (Taton *et al.,* 2006). Following the successful construction of a library, fingerprinting techniques can be employed to distinguish between clones for sequencing and identification. This provides a phylogenetic snap-shot of the community under investigation. Nübel et al. (1997) designed a set of cyanobacterial-specific DGGE primers for high-throughput fingerprinting. The primers are designed to amplify both filamentous and unicellular types of cyanobacterial 16S rRNA genes. This strategy gives rapid information of cyanobacterial communities in a given environment.



Recent molecular based studies investigating cyanobacterial diversity in Antarctic soils have found contradicting results. Niederberger et al. (2007) found cyanobacterial signals only in samples with significant water availability.

However, Smith et al. (2006) identified numerous cyanobacterial species in a low moisture (0.6 wt. % $H_2O$) soil sample.

Here we report the findings of a phylogenetic survey on cyanobacteria from an Antarctic Miers Valley soil transect using DGGE analysis and 16S rDNA library construction.

### 3.2 Site description

Samples for this study were collected from the northern slope of the Miers Valley, one of the coastal Dry Valleys (78°6'S, 164°0'E) in Eastern Antarctica. Characteristics of all sites along the transect are listed in Table 3.1. Eleven samples from the transect were individually selected for analysis. Soil temperatures ranged from -0.5°C to 1.1°C with minor fluctuations between the sites. All soil samples were of gravel texture. Soil samples were labeled according to the altitude a.s.l. Sample 171 consisted of wet gravels from the flood plain near Lake Miers. Relative humidity (RH) measurements of the soils indicate changes along the transect. Sample 171 could not be measured due to water saturation caused by Lake Miers. Sample 184 produced the lowest RH reading (42%) and sample 634 the highest reading (97%).

**Table 3.1: Characteristics of the Miers Valley soil transect**

| Transect sample no. / altitude a.s.l (m) | Description of soil | Soil surface temperature (°C) | % Soil relative humidity |
|---|---|---|---|
| 171 | Wet gravel | -0.2 | Water saturated |
| 184 | Course gravel | -0.5 | 42 |
| 234 | Sorted gravel | -0.3 | 57 |
| 284 | Gravel | -0.4 | 79 |
| 334 | Fine gravel | -0.4 | 76 |
| 384 | Fine gravel | -0.2 | 92 |
| 434 | Fine gravel | -0.4 | 81 |
| 484 | Fine gravel | -0.2 | 93 |
| 534 | Fine gravel | +0.8 | 96 |
| 584 | Fine gravel | +1.1 | 97 |
| 634 | Fine gravel | -0.3 | 97 |

**-Soil relative humidity (%) values are extremely inaccurate above 85%**



### 3.3 Metagenomic DNA extraction

Metagenomic DNA was extracted as described in Section 2.7.1, in duplicate from all transect samples, and the DNA fractions pooled. The DNA yields obtained are listed in Table 3.2. Soil collected at the University of the Western Cape (Cape Town, South Africa) was included in the extraction as a positive control. Sterile quartz sand was used as a negative control.

High molecular weight DNA was obtained from all transect samples and the positive control sample (Fig. 3.1). The negative control sample yielded no visible DNA (Fig. 3.1, lane 15). Sample 171 showed a very weak signal in comparison to other samples with samples 434 and 634 produced the strongest bands after electrophoresis. No co-extracted RNA or humic acids were observed for the Antarctic transect samples, the positive control sample however, showed signs of RNA (Fig. 3.1, lane 14).


**Figure 3.1: Agarose gel (0.7%) showing metagenomic DNA extracted from the Miers Valley transect soil samples.** Lane 1 and 13, DNA molecular marker (In-house prep); Lane 2, sample 171; Lane 3, sample 184; Lane 4, sample 234; Lane 5, sample 284; Lane 6, sample 334; Lane 7, sample 384; Lane 8, sample 434; Lane 9, sample 484; Lane 10, sample 534; Lane 11, sample 584; Lane 12, sample 634; Lane 14, crude positive control sample; Lane 15, negative control.

**Table 3.2: Concentrations and purity of metagenomic DNA isolated from the Miers Valley soil transect**

| Soil sample | DNA yield per g of soil (ng/g) | Purity | |
|---|---|---|---|
| | | $A_{280}/ A_{260}$ | $A_{260}/ A_{230}$ |
| 171 | 33 | 1.84 | 1.92 |
| 184 | 88 | 1.58 | 1.90 |
| 234 | 395 | 1.70 | 1.89 |
| 284 | 368 | 1.92 | 1.80 |
| 334 | 213 | 1.95 | 1.96 |
| 384 | 458 | 1.81 | 1.78 |
| 434 | 678 | 2.00 | 1.89 |
| 484 | 293 | 1.59 | 1.98 |
| 534 | 303 | 1.65 | 1.76 |
| 584 | 530 | 2.00 | 1.87 |
| 634 | 1068 | 1.88 | 1.94 |
| ¥Positive control | 1783 | 1.23 | 1.40 |

¥ Sample was subjected to a 1: 1000 dilution before measurements

Samples 434, 634 and the positive control yielded the highest concentrations of DNA. All Antarctic transect samples were of good quality with low concentrations of protein and humic acid co-extracted, as indicated by the $A_{280}/ A_{260}$ and $A_{260}/ A_{230}$ ratios. The positive control showed protein and humic acid contamination.



### 3.4 Cyanobacterial community analysis

#### 3.4.1 Homogeneity test using DGGE

DGGE analysis was performed to test for homogeneity of an altitudinal sample area. The sample site at 184 m a.s.l was selected for the test and metagenomic DNA was extracted from all samples (Fig. 3.2 [A]). PCR amplification as described in Section 2.13.1 was used with the extracted DNA as templates to prepare the DGGE amplicons.

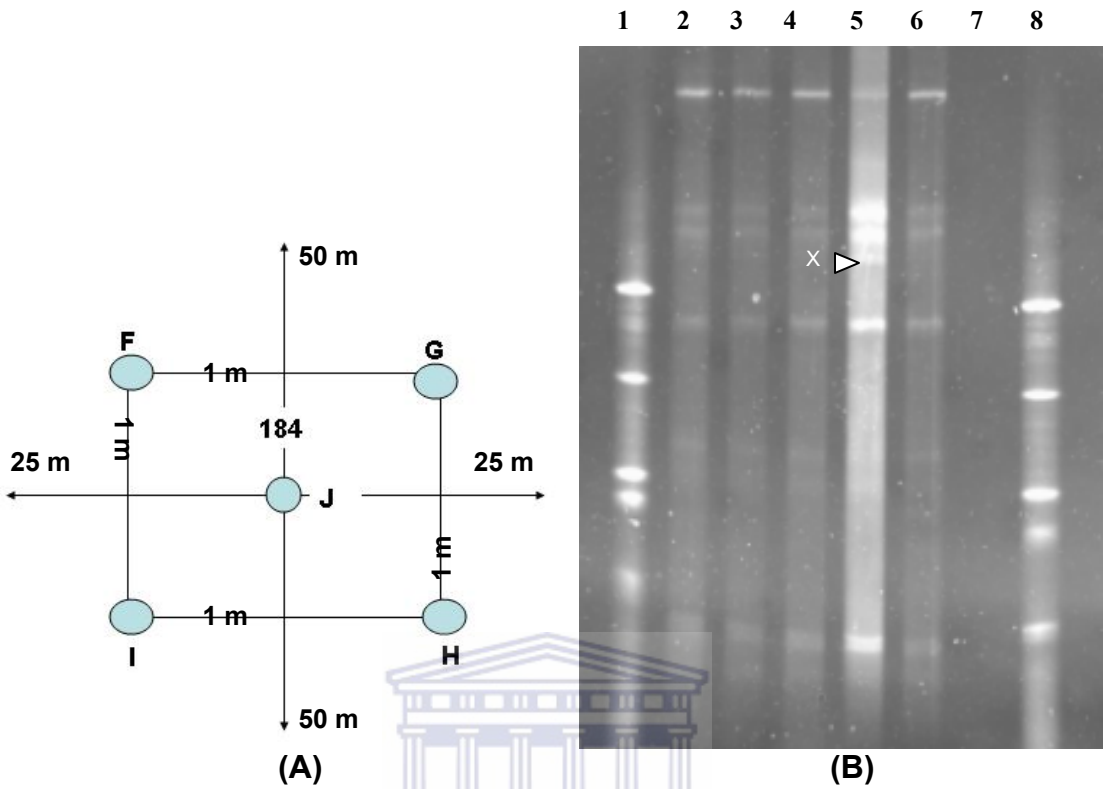


**Figure 3.2: Homogeneity test.** (A) Graphical view of the 184 subsamples. Pooled sample is indicated by J. (B) DGGE profile of replicates of the 184 sample site. Lanes 1 and 8, DGGE marker (In-house prep); Lane 2, 184F sample; Lane 3, 184G sample; Lane 4, 184H sample; Lane 5, 184I sample; Lane 6, 184J sample; Lane 7, Negative control.

The DGGE profile showed similar banding patterns for all the 184 samples. However, band type X (Fig. 3.2 [B], lane 5) was found more intense in the 184I sample.

Cluster analysis as described in Section 2.13.2.1 of the gel in Fig. 3.2 (B) confirms the homogeneity of the 184 sample site despite the different cluster grouping for sample 184I as indicated by the dendrogram (Fig. 3.3). This different grouping of the 184I sample is primarily due to the intense band type X. Due to

homogeneity of the 184 site, only one subsample was analysed per altitude throughout the study.

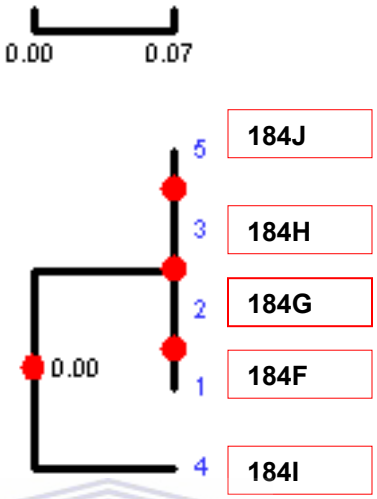


**Figure 3.3: Dendrogram constructed using Fig 3.2 as input using the similarity matrix based on the Dice coefficient and the UPGMA cluster method.** Samples from figure 3.2 are indicated in boxes.



### 3.4.2 Comparison of the cyanobacterial community structure along the Miers Valley transect

The metagenomic DNA extracts listed in Table 3.2 were used in a DGGE analysis to investigate the cyanobacterial diversity along the Miers Valley transect.

#### 3.4.2.1 Cyanobacterial 16S rDNA amplification for DGGE

The cyanobacterial PCR-DGGE strategy as described in Section 2.13.1 was used to prepare amplicons for DGGE analysis. This strategy was employed to overcome the problem of too little DNA for use as PCR templates. A PCR product of approximately 490 bp was expected after the second round PCR.

Amplicons of this size were obtained from all transect samples (Fig. 3.4) for further analysis by DGGE.

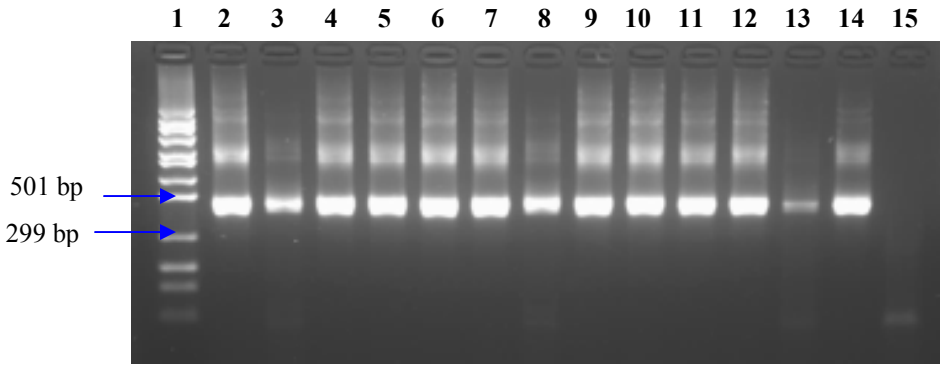


**Figure 3.4: Agarose gel (1.5%) showing amplicons obtained following 16S rDNA amplification using primers CYA359(T)F GC with CYA781aR/ CYA781bR on metagenomic DNA from all transect samples as templates.** Lanes 1, DNA marker (In-house prep); Lane 2, sample 171; Lane 3, sample 184; Lane 4, sample 234; Lane 5, sample 284; Lane 6, sample 334; Lane 7, sample 384; Lane 8, sample 434; Lane 9, sample 484; Lane 10, sample 534; Lane 11, sample 584, Lane 12, sample 634; Lane 13, Undiluted positive control sample used as PCR template; Lane 14, Diluted (1:1000) positive control sample used as PCR template; Lane 15, negative control.



### 3.4.2.2 Cyanobacterial DGGE fingerprints of the Miers Valley transect samples

Amplicons from Fig. 3.4 were separated by DGGE to analyze the cyanobacterial diversity across the Miers Valley transect. Fig. 3.5 shows the cyanobacterial DGGE profiles obtained from the transect samples. The most intense cyanobacterial band types are labeled on the gel and listed in Table 3.3.

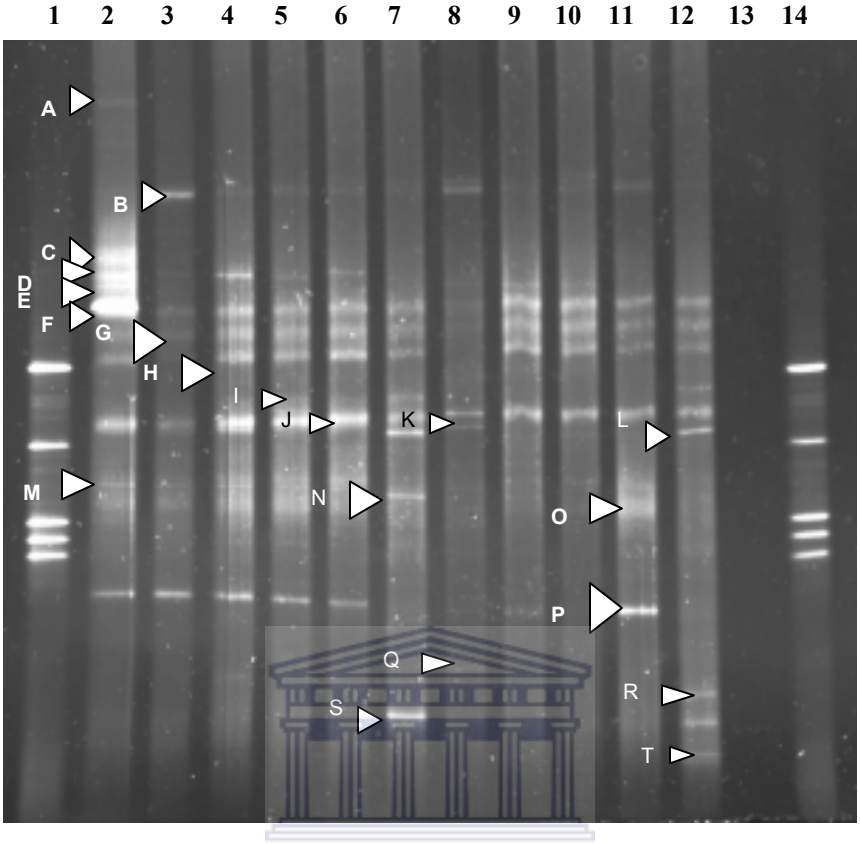


**Figure 3.5: DGGE profile of Miers Valley soil transect samples.** Lanes 1 and 14, DGGE marker (In-house prep); Lane 2, sample 171; Lane 3, sample 184; Lane 4, sample 234; Lane 5, sample 284; Lane 6, sample 334; Lane 7, sample 384; Lane 8, sample 434; Lane 9, sample 484; Lane 10, sample 534; Lane 11, sample 584; Lane 12, sample 634; Lane 13, Negative control. DGGE band types are indicated with [ ►] and labeled alphabetically.

These results show that there were differences in band types and community structure between the transect samples (Fig. 3.5). Samples 171, 384, 434 and 634 showed unique band type composition with higher cyanobacterial diversity than the other sites whereas the remaining samples gave virtually identical profiles.

**Table 3.3: Absence (-) and presence (+) of selected DGGE band types (Fig. 3.5) across the Miers Valley transect**

| DGGE band type | Miers valley transect sample no. | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 171 | 184 | 234 | 284 | 334 | 384 | 434 | 484 | 534 | 584 | 634 |
| A | + | - | - | - | - | - | - | - | - | - | - |
| B | - | + | + | + | + | + | + | + | + | + | + |
| C | + | - | - | - | - | - | - | - | - | - | - |
| D | + | + | + | + | + | - | - | - | - | - | - |
| E | + | - | - | - | - | - | + | + | + | - | - |
| F | + | + | + | + | + | + | + | + | + | + | + |
| G | + | + | + | + | + | + | + | + | + | + | + |
| H | + | + | + | + | + | + | + | + | + | + | + |
| I | - | - | - | + | + | + | - | - | - | - | + |
| J | + | + | + | + | + | + | + | + | + | + | + |
| K | - | - | - | - | - | - | + | - | - | - | - |
| L | - | - | - | - | - | + | - | - | + | - | + |
| M | + | + | + | + | + | - | + | + | + | + | - |
| N | - | - | - | - | - | + | - | - | - | - | - |
| O | + | + | + | + | + | - | - | - | - | + | - |
| P | + | + | + | + | + | - | - | + | - | + | - |
| Q | - | - | - | - | - | - | + | - | - | - | - |
| R | - | - | - | - | - | - | - | - | - | - | + |
| S | - | - | - | - | - | + | + | - | - | - | + |
| T | - | - | - | - | - | - | - | - | - | - | + |
| *20 | 11 | 9 | 9 | 10 | 10 | 9 | 10 | 8 | 8 | 8 | 10 |

* Indicates the total number of major band types



Twenty cyanobacterial band types were selected from Fig. 3.5 (Table 3.3). The highest number of bands was observed in the sample 171 followed by samples 284, 334, 434 and 634. Band types A and C were only found in the sample 171. Band type N was unique to sample 384. Band types K and Q were only found present in sample 434. Band types R and T were unique to sample 634. Interestingly, band type B occurred in all samples except for sample 171. Band types D, O and P were restricted to the lower parts of the transect, and band types F, G, H and J were the only band types present in all samples.

The banding patterns of the DGGE gel were further analysed by cluster analysis as described in Section 2.13.2.1 to visualize the differences between the transect samples. A dendrogram was constructed using Fig 3.5 as input. Three main cluster groups were observed clearly separating the lower altitude samples from the top altitude samples (Fig. 3.6). However, the top altitude samples were separated into two clusters (1 and 3). Interestingly, sample 534 (Fig. 3.5, lane 10) could not be clustered showing a unique DGGE pattern.

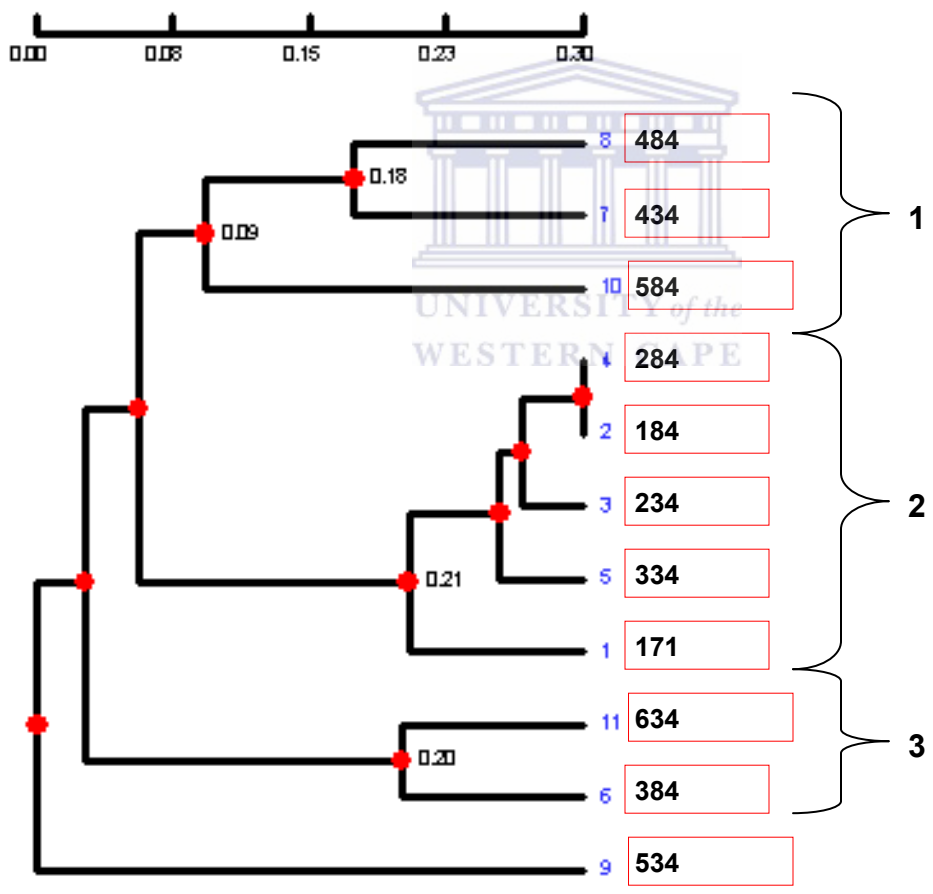


**Figure 3.6: Dendrogram constructed using Fig 3.5 as input using the similarity matrix based on the Dice coefficient and the UPGMA cluster method.** Samples are indicated in boxes. Clusters are indicated with [ **}** ].

### 3.5 Construction of cyanobacterial 16S rDNA libraries

To assess and compare the cyanobacterial diversity, four samples were selected. Samples 171, 184, 384 and 634 were selected based on their unique DGGE profiles and the PCR strategy described in Section 2.11.2 was followed.

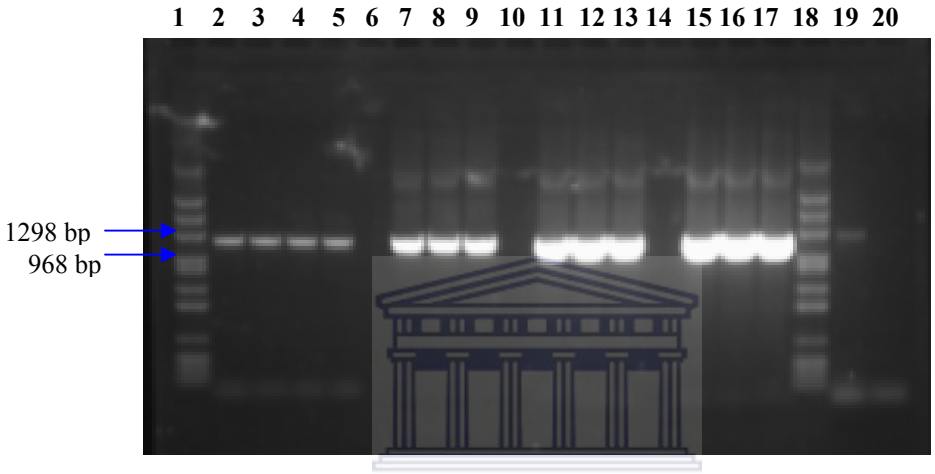


**Figure 3.7: Agarose gel (0.7%) showing amplicons obtained following 16S rDNA amplification using primer pair CYA359(T)F and U1510r.** Lane 1, DNA marker (In-house prep); Lanes 2, 3, 4, and 5, Sample of the PCR reaction using 171 as template; Lanes 7, 8, and 9, Sample of the PCR reaction using 184 as template; Lanes 11, 12, and 13, Sample of the PCR reaction using 384 as template; Lanes 15, 16, and 17, Sample of the PCR reaction using 634 as template; Lane 18, DNA marker (In-house prep); Lane 19, Sample of the PCR reaction using positive control as template; Lane 20, Sample of the PCR reaction using negative control as template.

Amplicons of the correct size (1.2 kb) were obtained from all four samples (Fig. 3.7). Amplicons were excised and purified from the agarose gel and ligated into the pTZ57R/T vector using the InsT/Aclone Kit (Fermentas) and transformed into electrocompetent *E. coli* DH10B cells.

### 3.6 Screening of recombinant clones

Screening of recombinant clones from the four libraries was done as a two step process: (1) using vector-derived M13 primers to identify clones with the correct insert size and (2) using the positive M13 products to identify cyanobacterial inserts using the primer pair CYA359(T)F and CYA781aR/ CYA781bR. Figure 3.8 shows colony PCR screening of 96 clones from the 384 library.

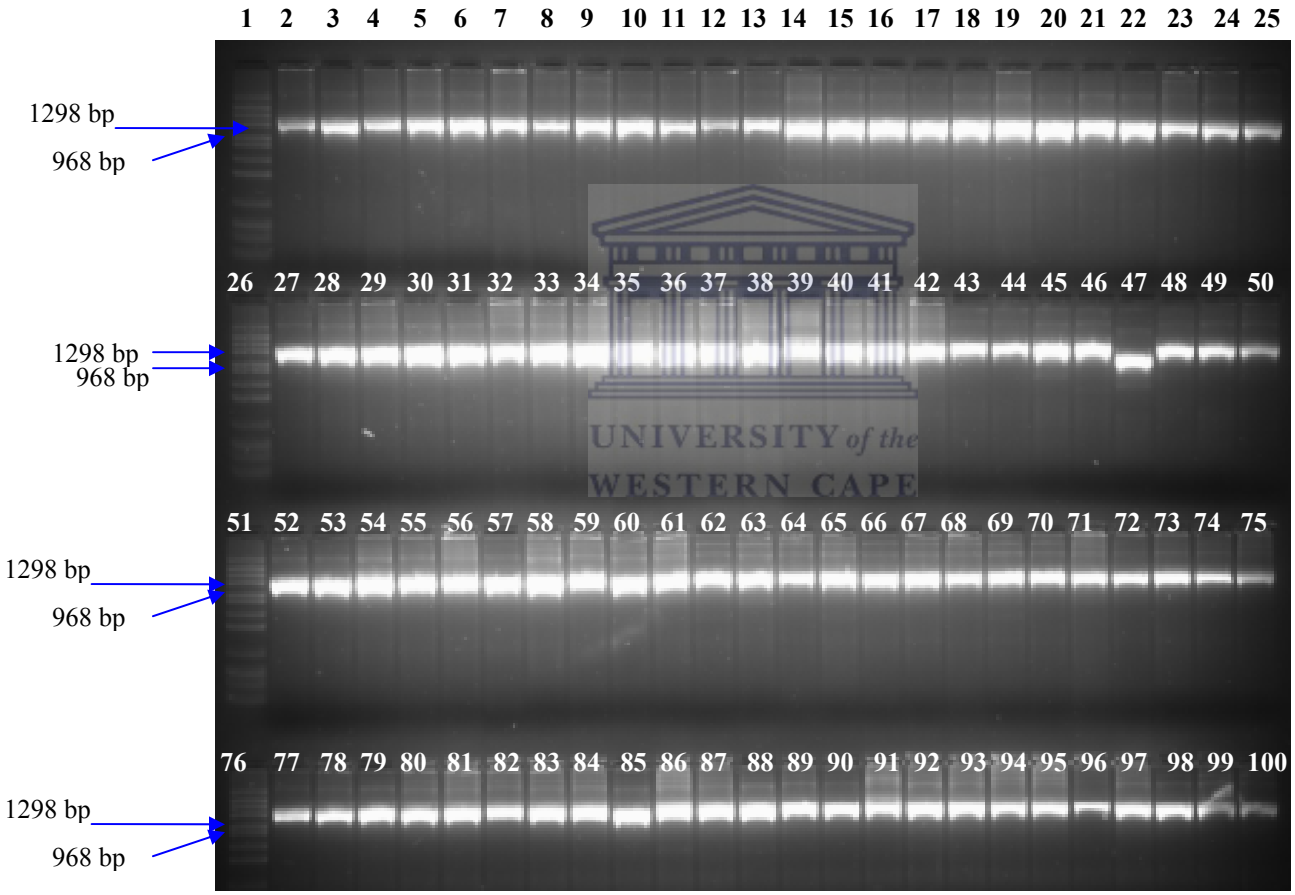


**Figure 3.8: Agarose gel (1%) showing amplicons obtained following colony PCR amplification using primers M13fw and M13rev on 96 positive clones from library 384.** Lanes 1, 26, 38, 51, 76, DNA molecular markers (In-house prep).

In total, 384 clones were analysed, 96 from each of the four libraries. All clones selected for screening with colony PCR contained inserts of the expected size.

The M13 PCR products were used as templates to screen for cyanobacterial inserts using primers CYA359(T)F and CYA781aR/ CYA781bR (Fig. 3.9 shows screening of library 384). A total of 232 putative cyanobacterial clones with a size of ~ 500 bps were obtained. In library 171, 50 cyanobacterial inserts were observed, 55 in library 184, 50 in library 384 and 77 in library 634. This result indicates that the primer pair used (Section 2.10) in the construction of the library was not entirely specific for cyanobacterial 16S rRNA.

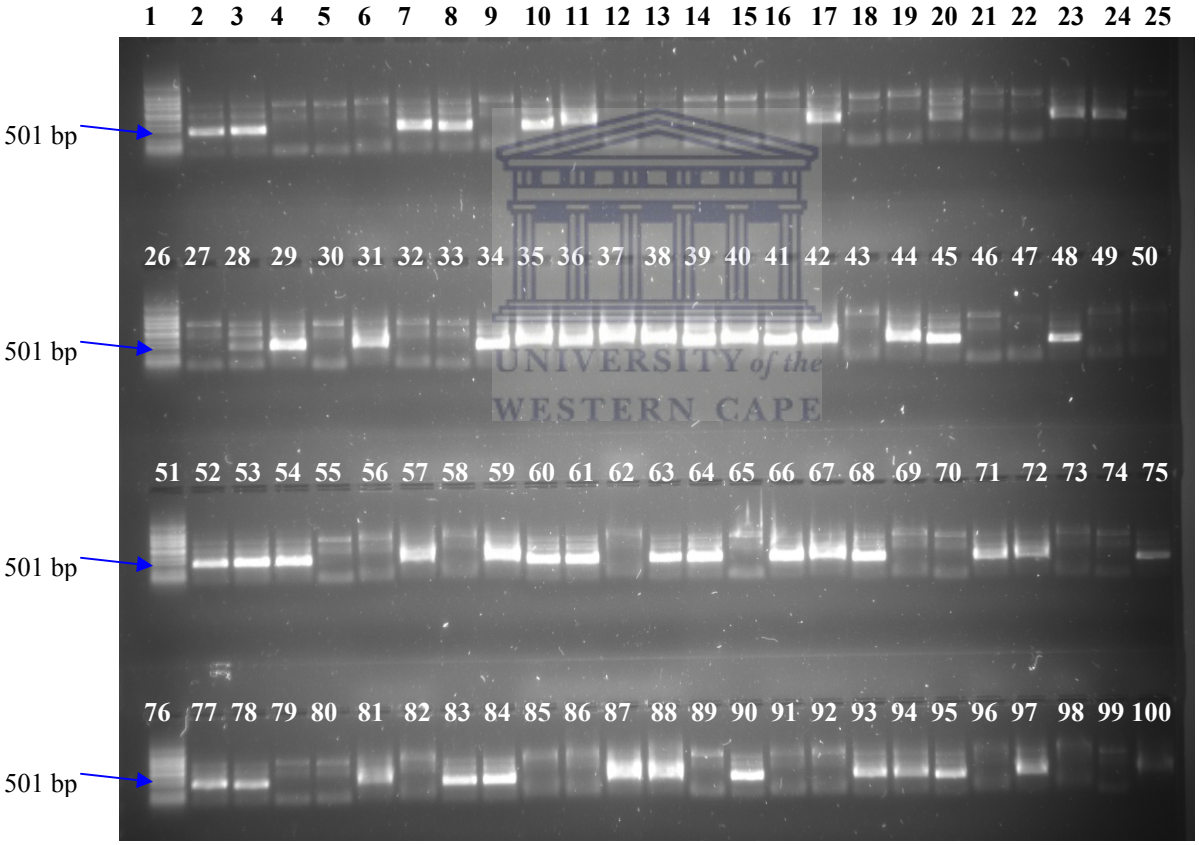


**Figure 3.9: Agarose gel (1.5%) showing amplicons obtained following PCR amplification using cyanobacterial specific primers (Section 2.11.5.1) on 96 positive clones from library 384.** Lanes 1, 26, 38, 51, 76, DNA molecular markers (In-house prep).

### 3.7 ARDRA analysis of putative cyanobacterial 16S rDNA inserts

The inserts of recombinant clones from the four libraries were reamplified using the primer pair CYA359(T)F and U1510r. A total number of 384 amplicons of 1.2 kb were retrieved and subjected to ARDRA using *Hae*III and *Rsa*I restriction enzymes in separate reactions. The reason for the inclusion of the non-cyanobacterial clones was the possibility of finding novel cyanobacterial 16S rDNA which elude PCR amplification using the primer pair CYA359(T)F and CYA781aR/ CYA781bR.

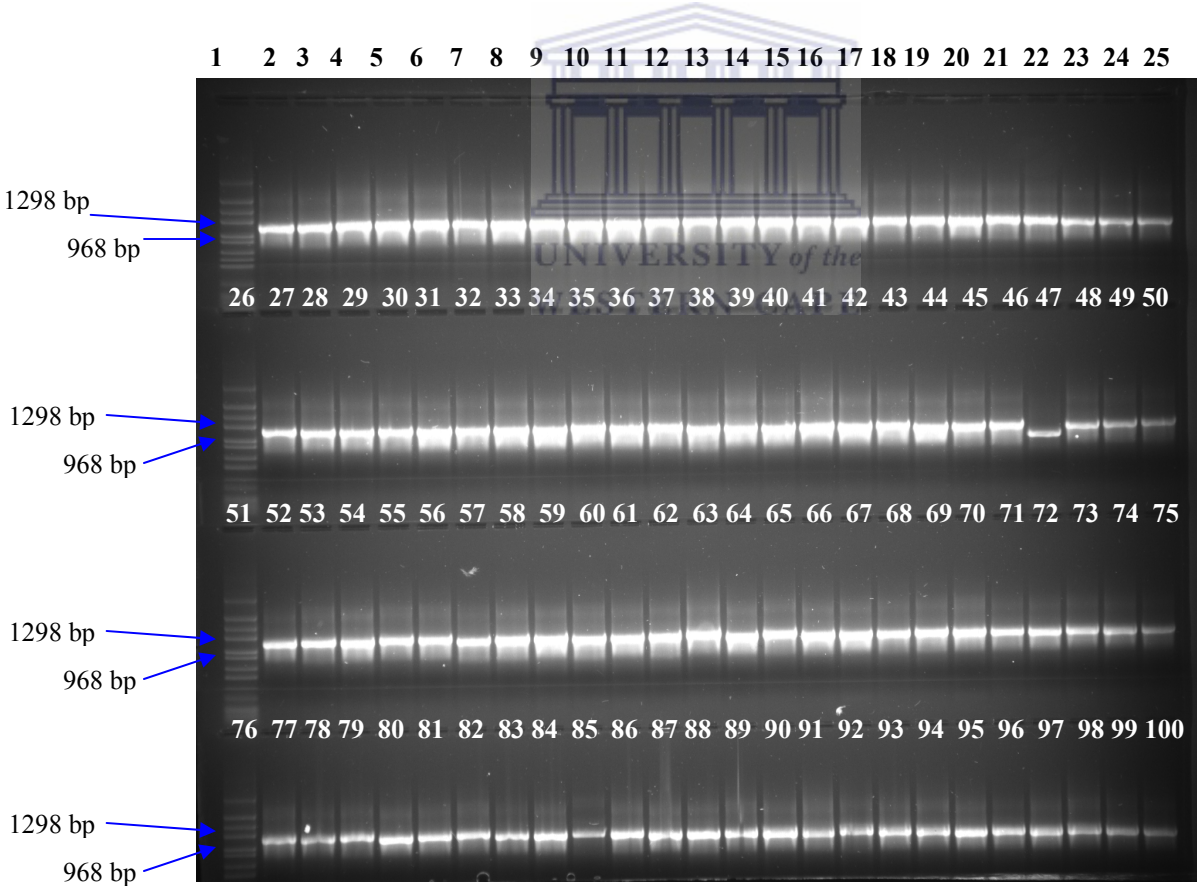


**Figure 3.10: Agarose gel (1%) showing amplicons obtained from library 384 for ARDRA analysis following PCR amplification as described in Section 2.11.6.** Lanes 1, 26, 38, 51, 76, DNA molecular markers (In-house prep).

The positive amplicons from Fig. 3.10 were subjected to ARDRA analysis. The obtained patterns from 48 clones from library 384 are shown in Figures 3.11 and 3.12.

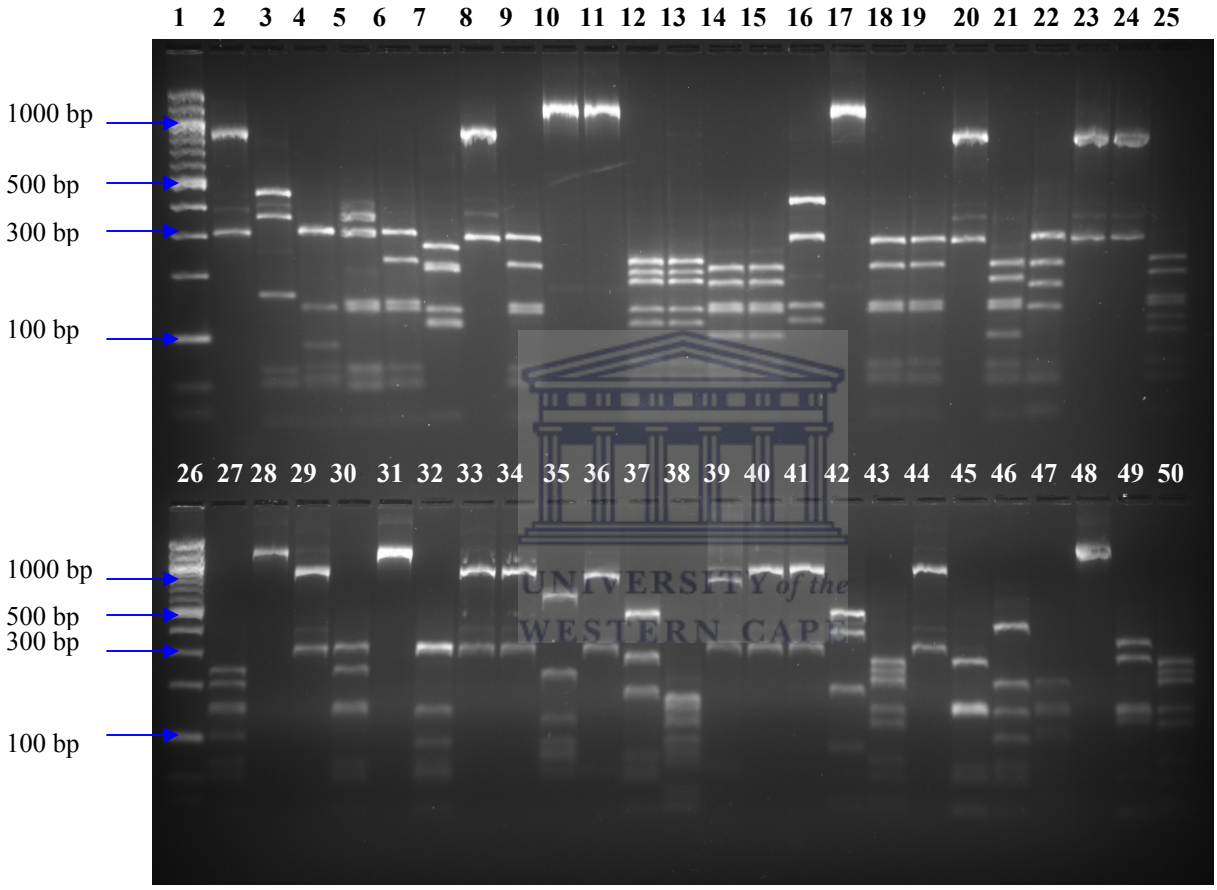


**Figure 3.11: Agarose gel (3%) showing ARDRA analysis of amplicons from library 384 with *Hae*III endonuclease.** Lanes 1 and 26, DNA molecular markers (Biolab 100 bp).

Restriction analysis was performed using the AlphaIEaseFC™ (version 4.0) software package (Alpha Innotech, San Leandro, CA). All restriction fragments were compared against the Biolab 100 bp marker and each band in the pattern assigned a molecular weight. All restriction patterns obtained with both enzymes

were compared for all libraries and ARDRA phylogroups were constructed. All putative cyanobacterial ARDRA groups (patterns obtained from clones that gave signals with semi-nested primers) were treated separately from non-cyanobacterial groups for further analysis.

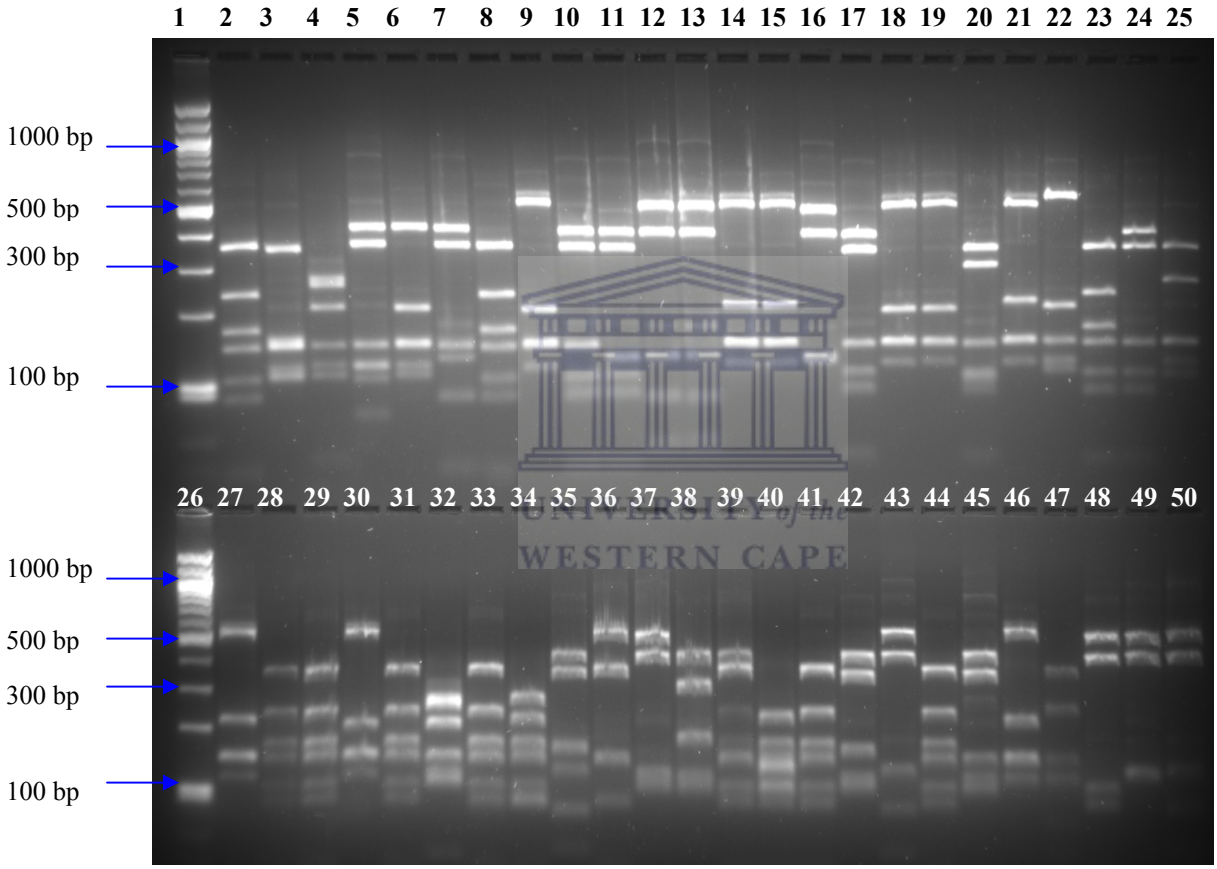


**Figure 3.12: Agarose gel (3%) showing amplicons from library 384 subjected to ARDRA analysis with *Rsa*I enzyme.** Lanes 1 and 26, DNA molecular markers (Biolab 100 bp).

A total of 18 distinct putative cyanobacterial ARDRA groups were retrieved (Table 3.4), covering 53.5% of the total cyanobacteria screened. ARDRA groups observed only twice in a specific library made up 11% of all fragments, with library 171 having the most doublets (24%). Single patterns contributed 35% of

the total cyanobacterial inserts, with more than 40% in libraries 171, 184 and 384 and 19.5% in library 634.

A cross-comparison of the ARDRA groups from all libraries is presented in Figure 3.13. ARDRA patterns occurring twice in a specific library were merged into one group (bar Y, Fig. 3.13), and all the single patterns were also grouped (bar X, Fig. 3.13). On average, 5 groups were observed per library. Only group 45 overlapped in libraries 171 and 184. Group 61 was found to be the most dominant ARDRA group in library 171 and group 48 the most dominant in library 184. Groups 94, 98, 107, 105 and 104 overlapped between libraries 384 and 634. Group 8 was the only ARDRA group unique for library 384, where group 94 (28% clones) was the most dominant. ARDRA groups 97 and 90 were unique to library 634 with group 98 being the most abundant group in the library. The highest diversity of ARDRA groups was found in libraries 171 and 634.

**Table 3.4: Total and percentages' of all cyanobacterial ARDRA phylotypes obtained from libraries, 171, 184, 384, and 634**

| | Library | | | | | Percentage | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Group nr. | 171 | 184 | 384 | 634 | Total | %171 | %184 | %384 | %634 | Total % |
| 104 | 0 | 0 | 1 | 1 | 2 | 0.00 | 0.00 | 2.00 | 1.30 | 0.86 |
| 42 | 3 | 0 | 0 | 0 | 3 | 6.00 | 0.00 | 0.00 | 0.00 | 1.29 |
| 49 | 0 | 3 | 0 | 0 | 3 | 0.00 | 5.45 | 0.00 | 0.00 | 1.29 |
| 52 | 3 | 0 | 0 | 0 | 3 | 6.00 | 0.00 | 0.00 | 0.00 | 1.29 |
| 72 | 3 | 0 | 0 | 0 | 3 | 6.00 | 0.00 | 0.00 | 0.00 | 1.29 |
| 8 | 0 | 0 | 4 | 0 | 4 | 0.00 | 0.00 | 8.00 | 0.00 | 1.72 |
| 45 | 2 | 2 | 0 | 0 | 4 | 4.00 | 3.64 | 0.00 | 0.00 | 1.72 |
| 53 | 0 | 4 | 0 | 0 | 4 | 0.00 | 7.27 | 0.00 | 0.00 | 1.72 |
| 59 | 0 | 4 | 0 | 0 | 4 | 0.00 | 7.27 | 0.00 | 0.00 | 1.72 |
| 90 | 0 | 0 | 0 | 4 | 4 | 0.00 | 0.00 | 0.00 | 5.19 | 1.72 |
| 106 | 0 | 0 | 2 | 2 | 4 | 0.00 | 0.00 | 4.00 | 2.60 | 1.72 |
| 61 | 5 | 0 | 0 | 0 | 5 | 10.00 | 0.00 | 0.00 | 0.00 | 2.16 |
| 92 | 0 | 5 | 0 | 0 | 5 | 0.00 | 9.09 | 0.00 | 0.00 | 2.16 |
| 97 | 0 | 0 | 0 | 10 | 10 | 0.00 | 0.00 | 0.00 | 12.99 | 4.31 |
| 48 | 0 | 11 | 0 | 0 | 11 | 0.00 | 20.00 | 0.00 | 0.00 | 4.74 |
| 107 | 0 | 0 | 2 | 10 | 12 | 0.00 | 0.00 | 4.00 | 12.99 | 5.17 |
| 98 | 0 | 0 | 2 | 17 | 19 | 0.00 | 0.00 | 4.00 | 22.08 | 8.19 |
| 94 | 0 | 0 | 14 | 10 | 24 | 0.00 | 0.00 | 28.00 | 12.99 | 10.34 |
| Y | 12 | 2 | 4 | 8 | 26 | 24.00 | 3.64 | 8.00 | 10.39 | 11.21 |
| X | 22 | 24 | 21 | 15 | 82 | 44.00 | 43.64 | 42.00 | 19.48 | 35.34 |
| Total | 50 | 55 | 50 | 77 | 232 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |

Y represents patterns that occurred twice
X represents single patterns

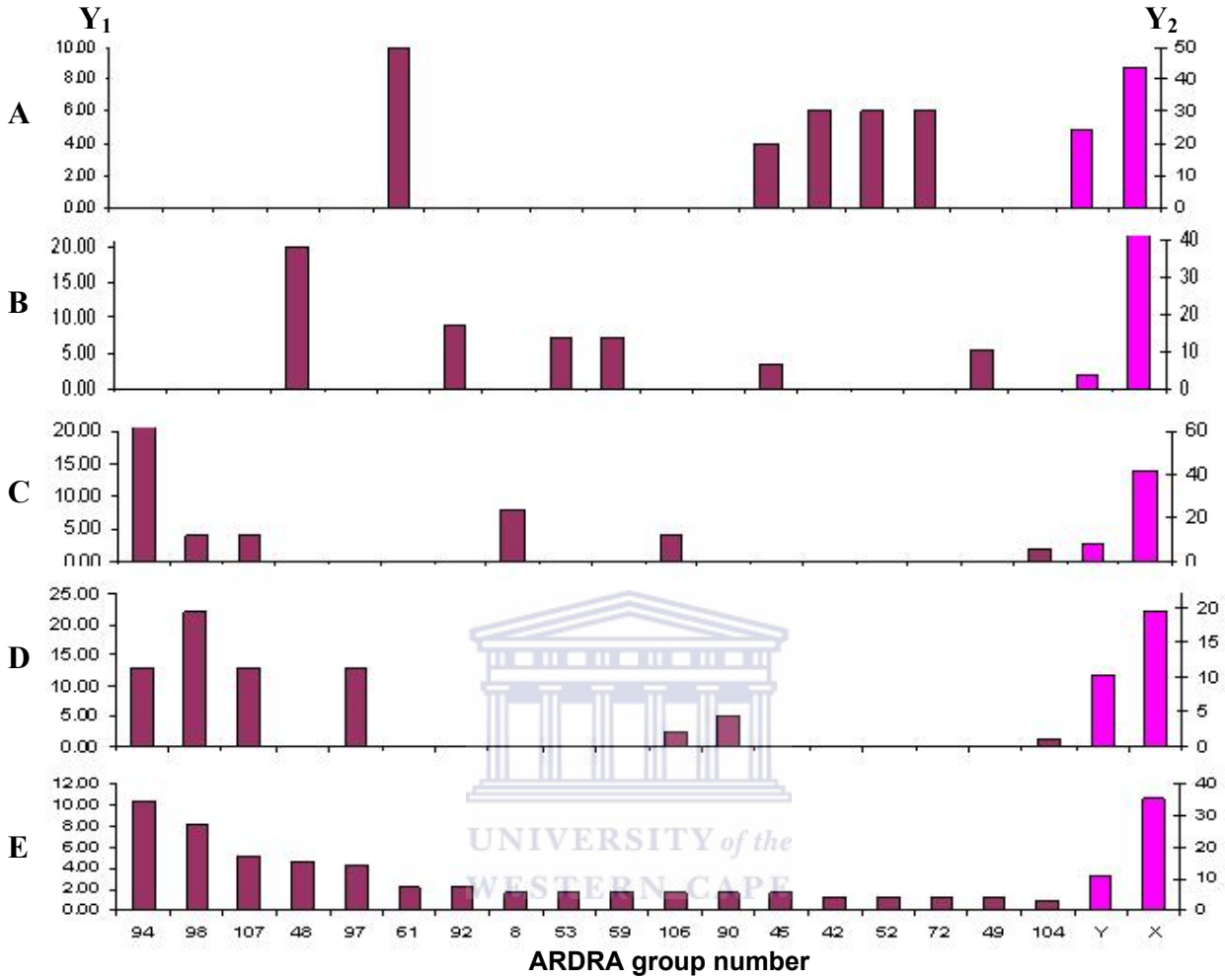


**Figure 3.13: Distribution of putative cyanobacterial ARDRA groups among the four 16S rDNA clone libraries of the Miers Valley transect.** (A) Library 171, (B) library 184, (C) library 384, (D) library 634, and (E) Total from all libraries. The percentage of clones present in ARDRA groups is represented by ■ and the values on the $Y_1$ axis. The percentage of clones present in ARDRA groups containing double (Y) and single patterns (X) is represented by ■ and the values on the $Y_2$ axis.

### 3.7.1 Richness estimation of ARDRA groups

To estimate the total number of putative cyanobacterial ARDRA groups from each library, the richness estimator $S_{Chao1}$ was calculated using the online software (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). It is calculated using the following formula:

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

Where $S_{obs}$ is the number of phylotypes observed in the library, and $F_1$ and $F_2$ are the number of phylotypes occurring as singletons or doubletons respectively.



Fig. 3.14 indicates the predicted number of ARDRA groups in the libraries based on $S_{chao1}$ estimator. The shape of the curves can be used as an indication of coverage of the libraries investigated. All the curves constructed indicated that the numbers of predicted ARDRA groups are much higher than the actual values obtained. An unexpected high diversity of putative cyanobacterial groups was predicted, which was much higher than values previously predicted for Antarctic soils.

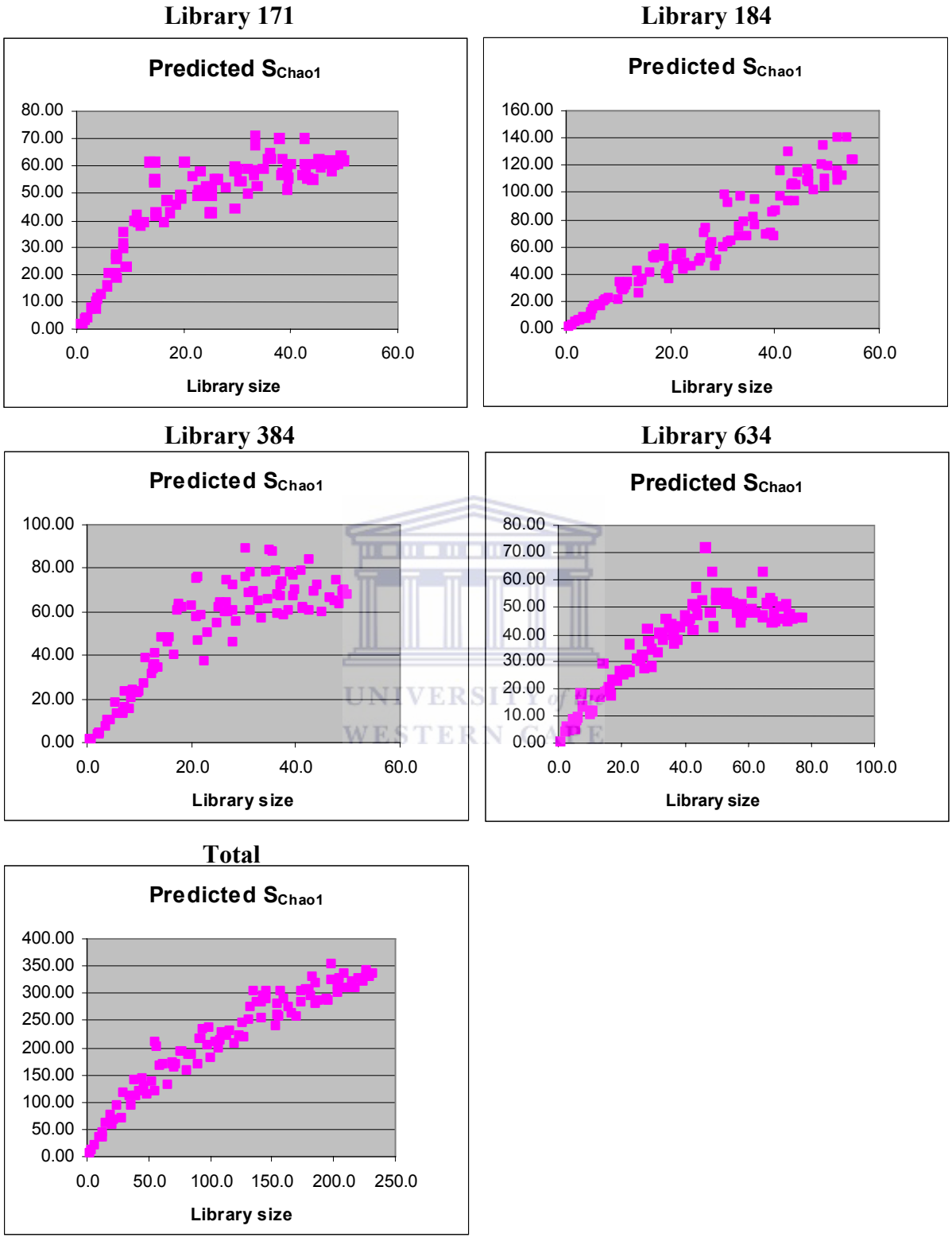


Figure 3.14: Predicted $S_{Chao1}$ results. Graphs show predicted number of ARDRA groups versus library size.

**Table 3.5: Summary of results obtained from the Evaluate library size online software**

| Library | Number of clones in library | Number of ARDRA groups observed | $S_{Chao1}$ (prediction for no. of ARDRA groups) |
|---|---|---|---|
| 171 | 50 | 33 | 62 |
| 184 | 55 | 31 | 123 |
| 384 | 50 | 29 | 68 |
| 634 | 77 | 26 | 46 |
| All libraries combined | 232 | 113 | 332 |

The Evaluate library size tool predicted more phylotypes in each library than the original values obtained by ARDRA (Table 3.5). Values of the $S_{chao1}$ estimator indicated a much higher number of ARDRA groups present in individual libraries as well as for the libraries combined. This result show that more clones need to be screened to reach saturation of cyanobacterial phylotypes in the libraries.



### 3.8 16S rDNA sequence analysis and distribution

Partial sequence information was obtained using the vector-derived M13 primers (Table 2.2) as described in Section 2.12.2. Approximately 50% of the clones were selected from each ARDRA group as listed in Table 3.4. A total of 101 clones were selected for sequencing and are listed in Table 3.6. The sequencing of seventeen of the clones failed after sequencing attempts (Table 3.6). A sequence length of approximately 850 bp was retrieved from the remaining 84 clones. Sequences were subjected to BLAST searches for the identification of the best matches in the GenBank database. The results are listed in Table 3.7.

**Table 3.6: Clones selected for sequencing from the constructed 16S rDNA libraries: (A) 171, (B) 184, (C) 384 and (D) 634**

| ARDRA group no. | Library | | | |
|---|---|---|---|---|
| | A | B | C | D |
| 94 | | | 7,22,38,83 | 5,11,20,**44** |
| 98 | | | 23,37 | 10,22,62,84 |
| 107 | | | 9,16 | 61,68,78 |
| 48 | | 9,10,**57**,**66**,83 | | |
| 97 | 33,**35**,45,75 | | | 57,**67**,82,**93** |
| 61 | | | | |
| 92 | | 4,39,45 | | |
| 8 | | | 13,20,62,66 | |
| 53 | | 2,40 | | |
| 59 | | 38,79 | | |
| 106 | | | 26,61 | **45** |
| 90 | | | | 7,43,53,63 |
| 45 | 2,82 | 13,**93** | | |
| 42 | 13,15 | | | |
| 52 | 50,**62** | | | |
| 72 | 8,**63** | | | |
| 49 | | 17,84 | | |
| 104 | | | **31** | 32 |
| Doubletons (Y) | 77,**14**,32,42,47,22 | **16** | 4,78 | 37,38,24,39 |
| Singletons (X) | 20,27,85 | 61,**6**,30 | 67,**44**,76 | 92,74,87 |
| Non-cyanos | 4,5,12 | 82,90,33 | 54,**14**,52 | **17**,88,72 |
| TOTAL | 24 | 23 | 23 | 31 |

Numbers indicates code numbers for each clone
Clones highlighted with red could not be sequenced

**Table 3.7: BLAST results of the sequenced clones listed in Table 3.6**

| Clone | Nearest match in GenBank | Site description of nearest match | Predicted phylogenetic group | Accession number | E-value | %I.D. |
|---|---|---|---|---|---|---|
| A33 | Uncultured cyanobacterium clone 500M2_B2 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 500m from glacier terminus replicate, Peru | Cyanobacteria | DQ514206 | 0.0 | 98% |
| A45 | Uncultured cyanobacterium clone FQSS027 16S ribosomal RNA gene, partial sequence | Rocky Mountain Endolithic sandstone community | Cyanobacteria | EF522247 | 0.0 | 93% |
| A75 | Uncultured cyanobacterium clone FQSS027 16S ribosomal RNA gene, partial sequence | Rocky Mountain endolithic sandstone community | Cyanobacteria | EF522247 | 0.0 | 94% |
| A2 | Uncultured cyanobacterium clone FQSS027 16S ribosomal RNA gene, partial sequence | Rocky Mountain endolithic sandstone community | Cyanobacteria | EF522247 | 0.0 | 94% |
| A82 | Limnothrix redekei CCAP 1443/1 16S rRNA gene (partial), ITS1, tRNA-Ile gene, tRNA-Ala gene and 23S rRNA gene (partial), strain CCAP 1443/1 | Shallow eutrophic lake | Cyanobacteria | AJ580007 | 0.0 | 97% |
| A13 | Uncultured Antarctic bacterium LB3-47 16S ribosomal RNA gene, partial sequence | Antarctic lake ice | Cyanobacteria | AF076163 | 0.0 | 98% |
| A15 | Uncultured Antarctic bacterium LB3-47 16S ribosomal RNA gene, partial sequence | Antarctic lake ice | Cyanobacteria | AF076163 | 0.0 | 98% |
| A50 | Uncultured cyanobacterium gene for 16S rRNA, partial sequence, clone: pltb-vmat-79 | Microbial mat at a shallow submarine hot spring, Taketomi Island, Japan | Cyanobacteria | AB294971 | 0.0 | 94% |
| A8 | Limnothrix redekei CCAP 1443/1 16S rRNA gene (partial), ITS1, tRNA-Ile gene, tRNA-Ala gene and 23S rRNA gene (partial), strain CCAP 1443/1 | Shallow eutrophic lake | Cyanobacteria | AJ580007 | 0.0 | 98% |

**Table 3.7: Continued**

| Clone | Nearest match in GenBank | Site description of nearest match | Predicted phylogenetic group | Accession number | E-value | %I.D. |
|---|---|---|---|---|---|---|
| A20 | Uncultured bacterium clone FFCH10450 16S ribosomal RNA gene, partial sequence | Soil from an undisturbed mixed grass prairie preserve | Uncultured organism | EU132325 | 0.0 | 93% |
| A27 | Uncultured bacterium gene for 16S rRNA, partial sequence, clone: BS137 | Bulk soil of reed bed reactor | Uncultured organism | AB240271 | 0.0 | 98% |
| A85 | Uncultured cyanobacterium clone RBE1CI-123 16S ribosomal RNA gene, partial sequence | Bogota River site E1-Se-SRp, Colombia | Cyanobacteria | EF111085 | 0.0 | 98% |
| A4 | Uncultured cyanobacterium clone FQSS027 16S ribosomal RNA gene, partial sequence | Rocky Mountain endolithic sandstone community | Cyanobacteria | EF522247 | 0.0 | 95% |
| A5 | Virgibacillus carmonensis partial 16S rRNA gene, type strain LMG 20964T | Deteriorated mural paintings | Firmicutes | AJ316302 | 0.0 | 99% |
| A12 | Uncultured Acidobacteria bacterium clone 351B 16S ribosomal RNA gene, partial sequence | Hydrocarbon-contaminated soil, Antarctica | Acidobacteria | AY571792 | 0.0 | 99% |
| A77 | Uncultured Acidobacteria bacterium clone AKYG1751 16S ribosomal RNA gene, partial sequence | Farm soil adjacent to a silage storage bunker, Minnesota, USA | Acidobacteria | AY921983 | 0.0 | 97% |
| A32 | Uncultured cyanobacterium clone RBE1CI-123 16S ribosomal RNA gene, partial sequence | Bogota River site E1-Se-SRp, Colombia | Cyanobacteria | EF111085 | 0.0 | 98% |
| A42 | Uncultured cyanobacterium clone RBE1CI-123 16S ribosomal RNA gene, partial sequence | Bogota River site E1-Se-SRp, Colombia | Cyanobacteria | EF111085 | 0.0 | 96% |
| A47 | Limnothrix redekei CCAP 1443/1 16S rRNA gene (partial), ITS1, tRNA-Ile gene, tRNA-Ala gene and 23S rRNA gene (partial), strain CCAP 1443/1 | Shallow eutrophic lake | Cyanobacteria | AJ580007 | 0.0 | 98% |

**Table 3.7: Continued**

| Clone | Nearest match in GenBank | Site description of nearest match | Predicted phylogenetic group | Accession number | E-value | %I.D. |
|---|---|---|---|---|---|---|
| A22 | Uncultured organism clone EME017 16S ribosomal RNA gene, partial sequence | Ice, Antarctica | Uncultured organism | EF127595 | 0.0 | 98% |
| | | | | | | |
| B45 | Uncultured organism clone EME118 16S ribosomal RNA gene, partial sequence | Ice, Antarctica | Uncultured organism | EF127594 | 0.0 | 99% |
| B4 | Uncultured organism clone EME017 16S ribosomal RNA gene, partial sequence | Ice, Antarctica | Uncultured organism | EF127595 | 0.0 | 96% |
| B39 | Uncultured Acidobacteria bacterium clone 351B 16S ribosomal RNA gene, partial sequence | Hydrocarbon-contaminated soil, Antarctica | Acidobacteria | AY571792 | 0.0 | 99% |
| B2 | Uncultured cyanobacterium clone RBE1CI-123 16S ribosomal RNA gene, partial sequence | Bogota River site E1-Se-SRp, Colombia | Cyanobacteria | EF111085 | 0.0 | 98% |
| B40 | Uncultured cyanobacterium clone RBE1CI-123 16S ribosomal RNA gene, partial sequence | Bogota River site E1-Se-SRp, Colombia | Cyanobacteria | EF111085 | 0.0 | 98% |
| B38 | Uncultured Antarctic bacterium LB3-47 16S ribosomal RNA gene, partial sequence | Antarctic lake ice | Cyanobacteria | AF076163 | 0.0 | 97% |
| B79 | Uncultured Antarctic bacterium LB3-47 16S ribosomal RNA gene, partial sequence | Antarctic lake ice | Cyanobacteria | AF076163 | 0.0 | 97% |
| B13 | Uncultured cyanobacterium clone FQSS027 16S ribosomal RNA gene, partial sequence | Rocky Mountain endolithic sandstone community | Cyanobacteria | EF522247 | 0.0 | 94% |
| B17 | Uncultured cyanobacterium clone RJ094 16S ribosomal RNA gene, complete sequence | Microbial mat from Lake Reid, Antarctica | Cyanobacteria | DQ181682 | 0.0 | 99% |
| B84 | Uncultured cyanobacterium clone RBE1CI-123 16S ribosomal RNA gene, partial sequence | Bogota River site E1-Se-SRp, Colombia | Cyanobacteria | EF111085 | 0.0 | 97% |
| B61 | Uncultured Rubrobacter sp. clone 354H 16S ribosomal RNA gene, partial sequence | Hydrocarbon-contaminated soil, Antarctica | Actinobacteria | AY571811 | 0.0 | 99% |

**Table 3.7: Continued**

| Clone | Nearest match in GenBank | Site description of nearest match | Predicted phylogenetic group | Accession number | E-value | %I.D. |
|---|---|---|---|---|---|---|
| B30 | Uncultured Acidobacteria bacterium clone CAL4 16S ribosomal RNA gene, partial sequence | Epilithic biofilms in Saint Callixtus Catacombs, Rome | Acidobacteria | DQ139453 | 0.0 | 96% |
| B82 | Glacier bacterium FJI10 16S ribosomal RNA gene, partial sequence | Subglacial sediments and ice | Proteobacteria | AY315180 | 0.0 | 98% |
| B90 | Uncultured bacterium clone FFCH10604 16S ribosomal RNA gene, partial sequence | Soil from an undisturbed mixed grass prairie preserve | Uncultured organism | EU132636 | 0.0 | 98% |
| B33 | Uncultured bacterium clone ERF-F3 16S ribosomal RNA gene, partial sequence | Metagenome of the Tinto River rhizosphere | Uncultured organism | DQ906076 | 0.0 | 95% |
| B9 | Uncultured cyanobacterium clone RBE1CI-123 16S ribosomal RNA gene, partial sequence | Bogota River site E1-Se-SRp, Colombia | Cyanobacteria | EF111085 | 0.0 | 98% |
| B10 | Uncultured cyanobacterium clone 0M1_D10 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 0m from glacier terminus replicate 1, Peru | Cyanobacteria | DQ513843 | 0.0 | 97% |
| B83 | Uncultured cyanobacterium clone RBE1CI-123 16S ribosomal RNA gene, partial sequence | Bogota River site E1-Se-SRp, Colombia | Cyanobacteria | EF111085 | 0.0 | 98% |
| | | | | | | |
| C31 | Uncultured organism clone EME118 16S ribosomal RNA gene, partial sequence | Ice, Antarctica | Uncultured organism | EF127594 | 0.0 | 98% |
| C9 | Uncultured Acidobacteria bacterium clone 352G 16S ribosomal RNA gene, partial sequence | Hydrocarbon-contaminated soil, Antarctica | Acidobacteria | AY571796 | 0.0 | 99% |
| C16 | Uncultured Acidobacteria bacterium clone 352G 16S ribosomal RNA gene, partial sequence | Hydrocarbon-contaminated soil, Antarctica | Acidobacteria | AY571796 | 0.0 | 99% |

**Table 3.7: Continued**

| Clone | Nearest match in GenBank | Site description of nearest match | Predicted phylogenetic group | Accession number | E-value | %I.D. |
|---|---|---|---|---|---|---|
| C26 | Uncultured Acidobacteria bacterium clone 352G 16S ribosomal RNA gene, partial sequence | Hydrocarbon-contaminated soil, Antarctica | Acidobacteria | AY571796 | 0.0 | 97% |
| C61 | Uncultured Acidobacteria bacterium clone 352G 16S ribosomal RNA gene, partial sequence | Hydrocarbon-contaminated soil, Antarctica | Acidobacteria | AY571796 | 0.0 | 97% |
| C23 | Uncultured soil bacterium clone M35_Pitesti 16S ribosomal RNA gene, complete sequence | Oil polluted soil from Romania | Uncultured organism | DQ378253 | 0.0 | 99% |
| C13 | Uncultured Pseudonocardia sp. clone 343G 16S ribosomal RNA gene, partial sequence | Hydrocarbon-contaminated soil, Antarctica | Actinobacteria | AY571815 | 0.0 | 99% |
| C20 | Uncultured Pseudonocardia sp. clone 343G 16S ribosomal RNA gene, partial sequence | Hydrocarbon-contaminated soil, Antarctica | Actinobacteria | AY571815 | 0.0 | 99% |
| C62 | Uncultured Pseudonocardia sp. clone 343G 16S ribosomal RNA gene, partial sequence | Hydrocarbon-contaminated soil, Antarctica | Actinobacteria | AY571815 | 0.0 | 99% |
| C66 | Uncultured Pseudonocardia sp. clone 343G 16S ribosomal RNA gene, partial sequence | Hydrocarbon-contaminated soil, Antarctica | Actinobacteria | AY571815 | 0.0 | 98% |
| C4 | Uncultured actinobacterium clone LV57-1-22 16S ribosomal RNA gene, partial sequence | Highly alkaline Lake Van, Turkey | Actinobacteria | AY642555 | 0.0 | 99% |
| C78 | Uncultured bacterium clone 101-85 16S ribosomal RNA gene, partial sequence | Heavy oil seeps of the Rancho La Brea tar pits | Uncultured organism | EF157177 | 0.0 | 89% |
| C54 | Uncultured bacterium clone FFCH3294 16S ribosomal RNA gene, partial sequence | Soil from an undisturbed mixed grass prairie preserve | Uncultured organism | EU134180 | 0.0 | 91% |

**Table 3.7: Continued**

| Clone | Nearest match in GenBank | Site description of nearest match | Predicted phylogenetic group | Accession number | E-value | %I.D. |
|---|---|---|---|---|---|---|
| C52 | Uncultured bacterium clone 101-85 16S ribosomal RNA gene, partial sequence | Heavy oil seeps of the Rancho La Brea tar pits | Uncultured organism | EF157177 | 0.0 | 90% |
| C67 | Uncultured bacterium clone YSK16S-50 16S ribosomal RNA gene, partial sequence | Acid mine drainage, Yinshan Mine, China | Uncultured organism | EF613013 | 0.0 | 95% |
| C76 | Uncultured bacterium clone FFCH18042 16S ribosomal RNA gene, partial sequence | Soil from an undisturbed mixed grass prairie preserve | Uncultured organism | EU135394 | 0.0 | 90% |
| C7 | Uncultured organism clone DLE075 16S ribosomal RNA gene, partial sequence | Ice, Antarctica | Uncultured organism | EF127618 | 0.0 | 97% |
| C22 | Uncultured Acidobacteria bacterium clone 351B 16S ribosomal RNA gene, partial sequence | Hydrocarbon-contaminated soil, Antarctica | Acidobacteria | AY571792 | 0.0 | 97% |
| C38 | Uncultured bacterium clone FBP241 16S ribosomal RNA gene, partial sequence | Lichen-dominated Antarctic cryptoendolithic community | Uncultured organism | AY250867 | 0.0 | 93% |
| C83 | Uncultured organism clone EME118 16S ribosomal RNA gene, partial sequence | Ice, Antarctica | Uncultured organism | EF127594 | 0.0 | 98% |
| | | | | | | |
| D57 | Uncultured Acidobacteria bacterium clone 500M2_D6 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 500m from glacier terminus replicate 2, Peru | Acidobacteria | DQ514186 | 0.0 | 97% |
| D82 | Uncultured Acidobacteria bacterium clone 500M2_D6 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 500m from glacier terminus replicate 2, Peru | Acidobacteria | DQ514186 | 0.0 | 98% |

**Table 3.7: Continued**

| Clone | Nearest match in GenBank | Site description of nearest match | Predicted phylogenetic group | Accession number | E-value | %I.D. |
|---|---|---|---|---|---|---|
| D32 | Uncultured Acidobacteria bacterium clone 500M1_C12 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 500m from glacier terminus replicate 1, Peru | Acidobacteria | DQ514153 | 0.0 | 89% |
| D7 | Uncultured bacterium clone High2_11 16S ribosomal RNA gene, partial sequence | Great Barrier Reef calcareous sediments, Australia | Uncultured organism | DQ256524 | 0.0 | 95% |
| D43 | Unidentified bacterium clone D09_ELL01 16S gene, partial sequence | Unvegetated soil environments at the Ellsworth Mountains, Antarctica | Uncultured organism | EF220093 | 0.0 | 99% |
| D53 | Uncultured bacterium clone High2_11 16S ribosomal RNA gene, partial sequence | Great Barrier Reef calcareous sediments, Australia | Uncultured organism | DQ256524 | 0.0 | 94% |
| D63 | Uncultured bacterium clone FBP241 16S ribosomal RNA gene, partial sequence | Lichen-dominated Antarctic cryptoendolithic community | Uncultured organism | AY250867 | 0.0 | 93% |
| D88 | Uncultured Acidobacteria bacterium clone 500M1_C12 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 500m from glacier terminus replicate 1, Peru | Acidobacteria | DQ514153 | 0.0 | 98% |
| D72 | Uncultured Acidobacteria bacterium clone 100M1_A9 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 500m from glacier terminus replicate 1, Peru | Acidobacteria | DQ513981 | 0.0 | 97% |
| D92 | Uncultured organism clone DLE037 16S ribosomal RNA gene, partial sequence | Ice, Antarctica | Uncultured organism | EF127609 | 0.0 | 99% |
| D74 | Uncultured Rubrobacter sp. clone 354H 16S ribosomal RNA gene, partial sequence | Hydrocarbon-contaminated soil, Antarctica | Actinobacteria | AY571811 | 0.0 | 99% |

**Table 3.7: Continued**

| Clone | Nearest match in GenBank | Site description of nearest match | Predicted phylogenetic group | Accession number | E-value | %I.D. |
|---|---|---|---|---|---|---|
| D87 | Uncultured bacterium clone 101-104 16S ribosomal RNA gene, partial sequence | Heavy oil seeps of the Rancho La Brea tar pits | Uncultured organism | EF157190 | 0.0 | 95% |
| D10 | Uncultured soil bacterium clone M35_Pitesti 16S ribosomal RNA gene, complete sequence | Oil polluted soil from Romania | Uncultured organism | DQ378253 | 0.0 | 98% |
| D22 | Uncultured Acidobacteria bacterium clone 500M2_F10 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 500m from glacier terminus replicate 2, Peru | Acidobacteria | DQ514188 | 0.0 | 97% |
| D62 | Uncultured Acidobacteria bacterium clone 500M2_F10 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 500m from glacier terminus replicate 2, Peru | Acidobacteria | DQ514188 | 0.0 | 99% |
| D84 | Uncultured soil bacterium clone M35_Pitesti 16S ribosomal RNA gene, complete sequence | Oil polluted soil from Romania | Uncultured organism | DQ378253 | 0.0 | 99% |
| D5 | Uncultured organism clone DLE075 16S ribosomal RNA gene, partial sequence | Ice, Antarctica | Uncultured organism | EF127618 | 0.0 | 99% |
| D11 | Unidentified bacterium clone D09_ELL01 16S gene, partial sequence | Unvegetated soil environments at the Ellsworth Mountains, Antarctica | Uncultured organism | EF220093 | 0.0 | 98% |
| D20 | Uncultured Acidobacteria bacterium clone 500M2_D6 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 500m from glacier terminus replicate 2, Peru | Acidobacteria | DQ514186 | 0.0 | 98% |
| D61 | Uncultured Acidobacteria bacterium clone 100M1_A9 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 100m from glacier terminus replicate 1, Peru | Acidobacteria | DQ513981 | 0.0 | 96% |

**Table 3.7: Continued**

| Clone | Nearest match in GenBank | Site description of nearest match | Predicted phylogenetic group | Accession number | E-value | %I.D. |
|---|---|---|---|---|---|---|
| D68 | Uncultured Acidobacteria bacterium clone 100M1_A9 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 100m from glacier terminus replicate 1, Peru | Acidobacteria | DQ513981 | 0.0 | 97% |
| D78 | Uncultured Acidobacteria bacterium clone 100M1_A9 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 100m from glacier terminus replicate 1, Peru | Acidobacteria | DQ513981 | 0.0 | 97% |
| D37 | Uncultured bacterium clone DRV-B011 16S ribosomal RNA gene, partial sequence | Rock Varnish in the Whipple Mountains, California | Uncultured organism | AY923081 | 0.0 | 94% |
| D38 | Uncultured bacterium clone FBP241 16S ribosomal RNA gene, partial sequence | Lichen-dominated Antarctic cryptoendolithic community | Uncultured organism | AY250867 | 0.0 | 95% |
| D24 | Uncultured Acidobacteria bacterium clone 100M1_A9 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 100m from glacier terminus replicate 1, Peru | Acidobacteria | DQ513981 | 0.0 | 97% |
| D39 | Uncultured Acidobacteria bacterium clone 500M2_F10 16S ribosomal RNA gene, partial sequence | Newly deglaciated soil 500m from glacier terminus replicate 2, Peru | Acidobacteria | DQ514188 | 0.0 | 97% |

The similarity of the 84 selected sequences with published data ranged from 89% to 99% (Table 3.7). Cyanobacteria were not the only bacterium phylum present in the Miers Valley transect libraries although specificity for the group was tested. Cyanobacteria were only observed in the lower altitude samples; 171 and 184. Other bacterium phyla include a large number of Uncultured organisms, Firmicutes, Acidobacteria, Actinobacteria and Proteobacteria. A high number of sequences (30%) showed relative low homology to published sequences (<97%), suggesting the presence of novel species or genera in the Miers Valley soils.

### 3.8.1 Taxonomic assignment of sequences

Due to the high number of unclassified sequences identified by the BLAST search (Table 3.7), all sequences were submitted for classification as described in Section 2.12.2. All sequences were successfully classified as bacteria as expected and are listed in Table 3.8. The percentages obtained from the RDP Classifier indicate the confidence level of the prediction.

**Table 3.8: Classifier results for sequenced clones**

| Clone | Predicted phylum (%) | Predicted family (%) | Predicted species (%) |
|---|---|---|---|
| A33 | Cyanobacteria[100%] | Family 4.1[96%] | *Anabaena*[40%] |
| A45 | Cyanobacteria[100%] | Family 4.1[100%] | *Anabaena*[32%] |
| A75 | Cyanobacteria[100%] | Family 4.1[100%] | *Anabaena*[45%] |
| A2 | Cyanobacteria[100%] | Family 4.1[99%] | *Anabaena*[47%] |
| A82 | Cyanobacteria[100%] | Family 2.2[35%] | *Chroococcidiopsis*[35%] |
| A13 | Cyanobacteria[100%] | Family 4.1[80%] | *Cylindrospermum*[46%] |
| A15 | Cyanobacteria[100%] | Family 2.2[44%] | *Chroococcidiopsis*[44%] |
| A50 | Cyanobacteria[100%] | Family 4.1[59%] | *Cylindrospermum*[26%] |
| A8 | Cyanobacteria[100%] | Family 4.1[73%] | *Cylindrospermum*[56%] |
| A20 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| A27 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp6*[100%] |
| A85 | Cyanobacteria[100%] | Family 4.1[60%] | *Cylindrospermum*[50%] |
| A4 | Cyanobacteria[100%] | Family 4.1[100%] | *Anabaena*[50%] |
| A5 | Firmicutes[100%] | Bacillaceae[100%] | *Virgibacillus*[100%] |
| A12 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| A77 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp6*[100%] |
| A32 | Cyanobacteria[100%] | Family 4.1[43%] | *Cylindrospermum*[34%] |
| A42 | Cyanobacteria[100%] | Family 4.1[59%] | *Cylindrospermum*[50%] |
| A47 | Cyanobacteria[100%] | Family 4.1[88%] | *Cylindrospermum*[55%] |
| A22 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| | | | |
| B45 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| B4 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| B39 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| B2 | Cyanobacteria[100%] | Family 4.1[62%] | *Cylindrospermum*[59%] |
| B40 | Cyanobacteria[100%] | Family 4.1[53%] | *Cylindrospermum*[43%] |
| B38 | Cyanobacteria[100%] | Family 2.2[39%] | *Chroococcidiopsis*[36%] |
| B79 | Cyanobacteria[100%] | Family 2.2[36%] | *Chroococcidiopsis*[39%] |
| B13 | Cyanobacteria[100%] | Family 4.1[100%] | *Anabaena*[68%] |
| B17 | Cyanobacteria[100%] | Family 4.1[43%] | *Cylindrospermum*[33%] |
| B84 | Cyanobacteria[100%] | Family 4.1[57%] | *Cylindrospermum*[51%] |
| B61 | Actinobacteria[100%] | Rubrobacteraceae[100%] | *Rubrobacter*[100%] |
| B30 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp6*[100%] |
| B82 | Proteobacteria[100%] | Oxalobacteraceae[99%] | *Herbaspirillum*[42%] |
| B90 | Actinobacteria[100%] | Micromonosporaceae[33%] | *Longispora*[31%] |
| B33 | Actinobacteria[100%] | Micromonosporaceae[47%] | *Longispora*[35%] |
| B9 | Cyanobacteria[100%] | Family 4.1[47%] | *Cylindrospermum*[43%] |
| B10 | Cyanobacteria[100%] | Family 3.1[69%] | *Oscillatoria*[62%] |
| B83 | Cyanobacteria[100%] | Family 4.1[61%] | *Cylindrospermum*[47%] |
| | | | |
| C31 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| C9 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| C16 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| C26 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| C61 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| C23 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| C13 | Actinobacteria[100%] | Pseudonocardiaceae[79%] | *Streptoalloteichus*[39%] |
| C20 | Actinobacteria[100%] | Pseudonocardiaceae[93%] | *Streptoalloteichus*[57%] |
| C62 | Actinobacteria[100%] | Pseudonocardiaceae[92%] | *Crossiella*[48%] |

**Table 3.8: Continued**

| Clone | Predicted phylum (%) | Predicted family (%) | Predicted species (%) |
|---|---|---|---|
| C66 | Actinobacteria[100%] | Pseudonocardiaceae[82%] | *Streptoalloteichus*[28%] |
| C4 | Actinobacteria[99%] | Acidimicrobiaceae[41%] | *Acidimicrobium*[41%] |
| C78 | Gemmatimonadetes[42%] | Gemmatimonadaceae[42%] | *Gemmatimonas*[42%] |
| C54 | Thermomicrobia[92%] | Thermomicrobiaceae[92%] | *Thermomicrobium*[92%] |
| C52 | Gemmatimonadetes[38%] | Gemmatimonadaceae[38%] | *Gemmatimonas*[38%] |
| C67 | Actinobacteria[100%] | Acidimicrobiaceae[48%] | *Acidimicrobium*[48%] |
| C76 | TM7[100%] | None | *TM7_genera_incertae_sedis*[100%] |
| C7 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| C22 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| C38 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| C83 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| | | | |
| D57 | Acidobacteria[99%] | Acidobacteriaceae[99%] | *Gp4*[99%] |
| D82 | Acidobacteria[100%] | Acidobacteriaceae[98%] | *Gp4*[100%] |
| D32 | Proteobacteria[61%] | Desulfohalobiaceae[8%] | *Desulfonatronovibrio*[8%] |
| D7 | Acidobacteria[91%] | Acidobacteriaceae[91%] | *Gp4*[90%] |
| D43 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| D53 | Acidobacteria[77%] | Acidobacteriaceae[77%] | *Gp4*[76%] |
| D63 | Acidobacteria[98%] | Acidobacteriaceae[98%] | *Gp4*[98%] |
| D88 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| D72 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| D92 | Actinobacteria[100%] | Micromonosporaceae[43%] | *Longispora*[41%] |
| D74 | Actinobacteria[100%] | Rubrobacteraceae[100%] | *Rubrobacter*[100%] |
| D87 | Thermomicrobia[40%] | Thermomicrobiaceae[40%] | *Thermomicrobium*[40%] |
| D10 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| D22 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| D62 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| D84 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| D5 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| D11 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| D20 | Acidobacteria[100%] | Acidobacteriaceae[100%] | *Gp4*[100%] |
| D61 | Acidobacteria[99%] | Acidobacteriaceae[99%] | *Gp4*[99%] |
| D68 | Acidobacteria[100%] | Acidobacteriaceae[98%] | *Gp4*[100%] |
| D78 | Acidobacteria[98%] | Acidobacteriaceae[100%] | *Gp4*[98%] |
| D37 | Actinobacteria[100%] | Rubrobacteraceae[100%] | *Rubrobacter*[100%] |
| D38 | Acidobacteria[97%] | Acidobacteriaceae[97%] | *Gp4*[97%] |
| D24 | Acidobacteria[99%] | Acidobacteriaceae[99%] | *Gp4*[99%] |
| D39 | Acidobacteria[99%] | Acidobacteriaceae[99%] | *Gp4*[99%] |

The results obtained from the RDP Classifier corresponded to the BLAST results on the bacteria phyla predicted. However, unclassified organisms were resolved to family and species level.

As observed in the BLAST results (Table 3.7), cyanobacteria were only present in the lower altitude libraries (171 and 184). Some clones could be classified into cyanobacterial families but no sequence give high confidence levels (≥80%) at the species level (Table 3.8).

Clones A45, A75, A2, A4 and B13 were the only sequences classified with 100% confidence. These sequences were classified into cyanobacterial family 4.1. Seventy-nine percent of the Cyanobacteria identified were of novel origin based on very low confidence levels obtained. *Anabaena* sp., *Chroococcidiopsis* sp., *Cylindrospermum* sp. and *Oscillatoria* sp. were the four cyanobacterial species identified with very low confidence. One *Oscillatoria*-like species was only present in sample 184. Five clones (C78, C52, D32, D53 and D87) could not be resolved at the phylum level (<80% confidence) suggesting novel placement within the bacterial phylogenetic tree. The mid-slope (384) and the high altitude (634) sample were dominated by acidobacterial sequences. Interestingly, acidobacterial sequences similar to the group VII (*gp6*) species were only present in the lower altitude samples (171 and 184), whereas the group IV (*gp4*) species were present in all libraries. Actinobacteria were also present in libraries 184, 384 and 634, but not in library 171. Astonishingly, two clones: C54 and D87 showed similarity to the thermophilic bacterial phylum Thermomicrobia. This finding was unexpected due to the preferred thermophilic nature of the Thermomicrobia species. The highly diverse but yet to be cultured group TM7 was represented by clone C76.

### 3.8.2 Community composition analysis

The taxonomic assignment of the obtained 16S rDNA sequences was determined using the RDP classifier in order to determine the distribution of the different phylogenetic groups among the libraries. Only clones with ≥80% confidence at the phylum level (Table 3.8) were placed into the respective phyla (Fig. 3.15 A and B).

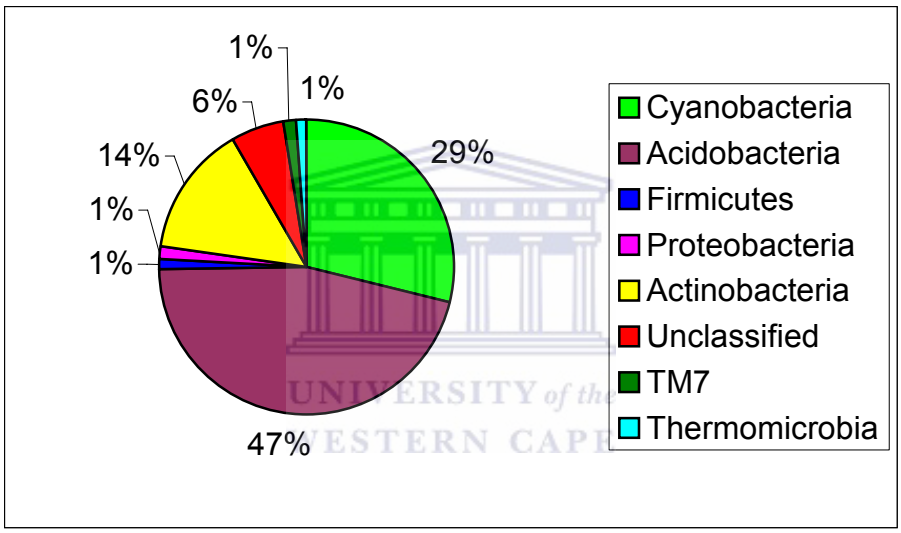


**Figure 3.15 A: Taxonomic assignment of the 16S rDNA sequences obtained from the combined libraries to different bacterial phyla.**

The clones sequenced from all libraries were used to compare the overall diversity between all samples based (Fig. 3.15 [A]). The phylum Acidobacteria is the overall dominant group containing 47% of all sequenced clones. The phylum Cyanobacteria, which are absent from libraries (samples) 384 and 634, make up 29% of the clones sequenced. Other groups present in the libraries include

Actinobacteria (14%), Unclassified sequences (6%), Firmicutes (1%), Proteobacteria (1%), TM7 (1%) and Thermomicrobia (1%).

The composition of the individual libraries (Fig. 3.15 [B]) shows that the phylum Cyanobacteria dominates the lower altitudinal samples (171 and 184), while Acidobacteria, which are present in all samples, dominate the mid-slope sample 384 and high altitude sample (634). Interestingly, unclassified sequences were only observed in the 384 and 634 samples. The phylum Actinobacteria was found to be absent in the sample 171 but present in all other libraries.

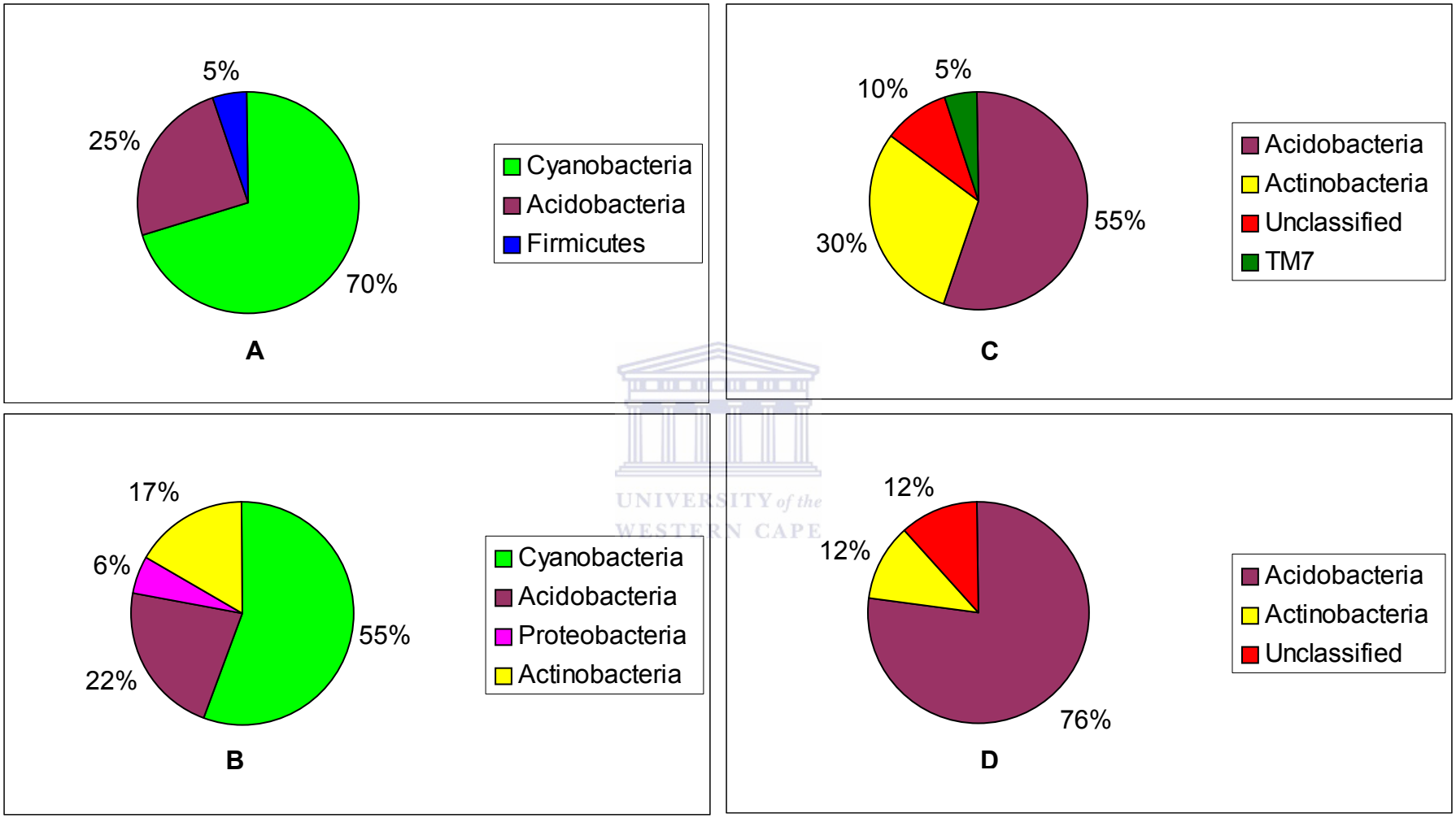


**Figure 3.15 B: Taxonomic assignment of the 16S rDNA sequences obtained from the libraries.** (A) 171, (B) 184, (C) 384 and (D) 634.

### 3.9 Investigating the non-specific binding of primer set CYA359(T)F and CYA781aR/ CYA781bR

The non-specific binding of the cyanobacterial primer set was investigated by alignment with the RDP Probe Match tool. The forward primer CYA359(T)F tended to bind to a large amount of cyanobacterial sequences in the RDP database at zero mismatches allowed (Table 4.1). However, when mismatches were introduced the specificity of the primer decreased drastically and other phyla were matched.

**Table 3.9: Probe Match alignment using primer CYA359(T)F.** Highlighted (yellow) regions indicates the phyla identified in this study

| Probe Match identity (Phylum) | Number of sequences matched per mismatche(s) allowed | | |
|---|---|---|---|
| | 0 | 1 | 2 |
| Chloroflexi | 1 | 7 | 67 |
| Cyanobacteria | 6883 | 7844 | 8035 |
| Proteobacteria | 22 | 3197 | 11284 |
| Firmicutes | 28 | 2163 | 13684 |
| Actinobacteria | 8 | 402 | 5315 |
| Acidobacteria | 4 | 4042 | 7230 |
| Bacteroidetes | 2 | 124 | 153 |
| Verrucomicrobia | 1 | 2 | 9 |
| Unclassified bacteria | 467 | 1484 | 4358 |
| Thermotogae | 0 | 1 | 2 |
| Nitrospira | 0 | 57 | 227 |
| Deferribacteres | 0 | 12 | 17 |
| Planctomycetes | 0 | 1 | 7 |
| Spirochaetes | 0 | 1 | 1 |
| TM7 | 0 | 7 | 159 |
| Aquificae | 0 | 0 | 5 |
| Thermodesulfobacteria | 0 | 0 | 2 |
| Deinococcus-Thermus | 0 | 0 | 1 |
| Fibrobacteres | 0 | 0 | 3 |
| Dictyoglomi | 0 | 0 | 7 |
| SR1 | 0 | 0 | 1 |
| Gemmatimonadetes | 0 | 0 | 186 |

The reverse primer CYA781aR was very specific at zero mismatches however, only 62 cyanobacterial sequences were matched in the database (Fig. 3.10). The primer specificity for cyanobacteria increased significantly as the mismatches for the primer increased but led to a much higher matching of other phyla.

**Table 3.10: Probe Match alignment using primer CYA781aR.** Highlighted (yellow) regions indicate the phyla identified in this study

| Probe match identity (Phylum) | Number of sequences matched per mismatche(s) allowed | | |
|---|---|---|---|
| | 0 | 1 | 2 |
| Cyanobacteria | 62 | 6885 | 7396 |
| Firmicutes | 1 | 118 | 5510 |
| Acidobacteria | 2 | 7 | 3554 |
| Verrucomicrobia | 1 | 1348 | 2303 |
| Unclassified bacteria | 6 | 534 | 4372 |
| Proteobacteria | 0 | 29 | 1970 |
| Planctomycetes | 0 | 6 | 2871 |
| Bacteroidetes | 0 | 2 | 479 |
| TM7 | 0 | 3 | 171 |
| OD1 | 0 | 80 | 105 |
| Actinobacteria | 0 | 0 | 81 |
| Thermotogae | 0 | 0 | 3 |
| Deinococcus-Thermus | 0 | 0 | 2 |
| Chloroflexi | 0 | 0 | 456 |
| Thermomicrobia | 0 | 0 | 8 |
| Nitrospira | 0 | 0 | 3 |
| Deferribacteres | 0 | 0 | 37 |
| Chlorobi | 0 | 0 | 2 |
| Spirochaetes | 0 | 0 | 14 |
| Fusobacteria | 0 | 0 | 15 |
| Dictyoglomi | 0 | 0 | 7 |
| Gemmatimonadetes | 0 | 0 | 4 |
| BRC1 | 0 | 0 | 36 |
| OP10 | 0 | 0 | 107 |
| OP11 | 0 | 0 | 45 |
| WS3 | 0 | 0 | 65 |
| SR1 | 0 | 0 | 16 |

Reverse primer CYA781bR was the most specific of the primers at zero mismatches however, only 21 cyanobacterial sequences were matched in the

database (Fig. 3.11). The primer did not align with any other phyla in the database at zero mismatches but showed non-specific binding at 1 and 2 mismatches.

**Table 3.11: Probe Match alignment using primer CYA781bR.** Highlighted (yellow) regions indicate the phyla identified in this study

| Probe match identity (Phylum) | Number of sequences matched per mismatche(s) allowed | | |
|---|---|---|---|
| | 0 | 1 | 2 |
| Cyanobacteria | 21 | 6806 | 7339 |
| Thermomicrobia | 0 | 1 | 8 |
| Proteobacteria | 0 | 26 | 1694 |
| Firmicutes | 0 | 117 | 3087 |
| Planctomycetes | 0 | 5 | 2834 |
| Acidobacteria | 0 | 4 | 3506 |
| Bacteroidetes | 0 | 2 | 385 |
| Verrucomicrobia | 0 | 1350 | 2288 |
| TM7 | 0 | 3 | 158 |
| OD1 | 0 | 80 | 103 |
| Unclassified bacteria | 0 | 517 | 4227 |
| Thermotogae | 0 | 0 | 3 |
| Deinococcus-Thermus | 0 | 0 | 2 |
| Chloroflexi | 0 | 0 | 448 |
| Nitrospira | 0 | 0 | 1 |
| Deferribacteres | 0 | 0 | 37 |
| Spirochaetes | 0 | 0 | 5 |
| Dictyoglomi | 0 | 0 | 7 |
| Gemmatimonadetes | 0 | 0 | 2 |
| BRC1 | 0 | 0 | 36 |
| OP10 | 0 | 0 | 105 |
| OP11 | 0 | 0 | 43 |
| WS3 | 0 | 0 | 64 |
| SR1 | 0 | 0 | 16 |



Interestingly, all non-cyanobacterial phyla identified in this study were matched when mismatches were introduced in the primers. However, the forward primer does not align to Thermomicrobia sequences. The forward primer tended to be

very sensitive to mismatches aligning to significant amounts of non-cyanobacterial phyla when 1 or 2 mismatches were introduced. The reverse primers were more specific at zero mismatches but did not align to a large set of cyanobacterial sequences. The reverse primers became more specific to Cyanobacteria when 1 or 2 mismatches were introduced.

# CHAPTER FOUR

## Discussion

### 4.1 Metagenomic DNA extraction and PCR amplification

Following sample collection, metagenomic DNA extraction is a fundamental step in microbial ecology. High molecular weight DNA is crucial in the analysis of the total microbial pool in any environment. Subsequent steps also depend on the purity of the DNA extracted. Due to the use of enzymes (e.g. DNA polymerases, restriction endonucleases, and T4 DNA ligases) for DNA manipulations, unpurified DNA contains substances that hinder the work of the enzyme. In Antarctic soils little organic matter persists due to complete absence of plant life especially in ice-free environments like the Miers Valley. This leads to extractions of pure DNA with no humic acid.



In this study high molecular weight DNA was obtained (Table 3.2 and Fig. 3.1) in the range of 33 ng to 1.1 µg per gram of soil from samples from a transect with increasing altitude in the Miers Valley, Eastern Antarctica. The lowest altitude sample (171 m a.s.l) gave the lowest DNA yield. This may be due to the wet gravel soil that provides protection against lysis or which binds naked DNA extremely tight. The DNA yields are difficult to compare to other findings due to limited studies on Antarctic soils. In a similar study Smith et al. (2006) reported values of 320 ng, 480 ng, and 840 ng per g soil for three different Miers Valley sites. Thus, results indicate that DNA yields in this study were in the same range as that obtained by Smith et al. (2006) for Miers Valley soils.

The extracted DNA was pure enough as expected (Table 3.2), and the subsequent experiments could be performed without extra purification steps. The DNA extracted was used for PCR amplification of a fragment of the cyanobacterial 16S rRNA gene. This was done successfully by using a nested-PCR strategy similar to that applied by Zwart et al. (2005).

### 4.2 Cyanobacterial DGGE fingerprinting

DGGE fingerprinting was conducted to analyse the cyanobacterial diversity across the Miers Valley soil transect. The procedure started with a homogeneity analysis of one of the transect sample site at 184 m a.s.l (Section 3.4.2). The result indicated that all sub-samples including the pooled sample 184J showed highly similar DGGE banding profiles (Fig 3.2). This allowed for the use of only the pooled transect samples for further cyanobacterial diversity analysis. The constructed DGGE profiles showed good resolution, and in total, 20 bands could be distinguished across the transect profiles (Table 3.3 and Fig. 3.5). Some of the DGGE bands were site-specific whereas others were common to all profiles. The DGGE profile across the transect showed some minor differences. Some samples produced virtually identical banding patterns. However, samples 171, 384, 434 and 634 showed a very different cyanobacterial community structure from the other samples. This result indicates that factors such as altitude and relative soil humidity (RH) might influence cyanobacterial community structure.

### 4.2.1 The effect of altitude and relative soil humidity on cyanobacterial community structure

The DGGE profile analyses used to estimate possible effects of factors such as altitude and relative soil humidity on the cyanobacterial community structure along the Miers Valley soil transect. The altitudinal variations indicated some restrictions on DGGE bands, most interestingly, band types B, C, O and P. Band type B was only absent from sample 171J whereas band types C, O and P were not present in the samples 384 m a.s.l and higher (Table 3.3). This may indicate that band type B may be restricted to dry soils, as sample 171J is predominantly sediment from Lake Miers, and band types C, O and P may be restricted to the lower altitude samples. This effect of the altitude on the cyanobacterial community structure may be due to soil water content, as the mid-slope samples show extremely low water content readings. Smith et al. (2006) reported water contents of 0.8% and 0.7% for samples 234 a high altitude sample (PENP), respectively, confirming the extremely low water content of the Miers Valley soils. Cyanobacteria are generally associated with moist regions (Miller and Bebout, 2004), making soil RH a key factor for their growth. Thus, differences of the cyanobacterial community structure within the transect might also be caused by differences in RH. The soil RH fluctuated from sample to sample, with site 171 not measured due to water saturation while samples 184, 234, 284 and 334 give readings under 90%. This was also observed for sample 434, which had a soil RH of 81%. The mid-slope sample 384, together with the top-slope samples 484, 534, 584 and 634 showed readings above 90%, the highest relative humidity

readings for the transect. The water-saturated site 171 showed the highest number of DGGE bands, which concurs with the statement of Miller and Bebout (2004). Interestingly, the DGGE dendrogram constructed from the DGGE gel (Fig. 3.5) shows that the lower altitude samples (171, 184, 234, 284 and 334) cluster together whereas the top altitude samples (384, 484, 534, 584 and 634) form a separate cluster. This result strongly suggests that the soil RH does shape the cyanobacterial community structure in the Miers Valley soil transect.

### 4.3 Cyanobacterial 16S rDNA libraries and ARDRA profiling

Cyanobacterial 16S rDNA clone libraries were constructed from sample sites 171, 184, 384 and 634. A total of 96 clones were selected for screening from each library. ARDRA analysis was used to discriminate between clones following successful screening. The resolution of the restricted amplicons on agarose gels was adequate for the classification of the clones into ARDRA groups. However, sequencing of specific clones from the groups revealed surprising results. Only some of the ARDRA groups were correctly resolved these include: group 8 (Actinobacteria), group 2 (Cyanobacteria), group 59 (Cyanobacteria) and group 42 (Cyanobacteria). All other groups did not deliver similar sequences as expected after ARDRA analysis. However, highly mixed sequences were observed within ARDRA groups. Probe Match analysis indicated that the cyanobacterial primer set were not entirely specific (Section 3.9) leading to broad range amplification of other bacterial genera.

Thus, ARDRA profiling suggested high cyanobacterial diversity for each individual sample site in contrast to the DGGE profiles. Richness estimation using putative ARDRA groups indicated that large numbers of cyanobacterial putative groups existed in the libraries. However, sequencing results disproved this conclusion by indicating other genera present in the libraries.

**4.4 Phylotype coverage and distribution**

A taxonomic assignment of the sequenced clones (Fig. 3.15) indicated that Acidobacteria (47%) were the most dominant in all the libraries analysed followed by Cyanobacteria (29%) and Actinobacteria (14%). Cyanobacteria dominated the lower altitude samples 171 and 184, but no cyanobacterial sequences were retrieved from samples 384 and 634. Niederberger et al. (2007) showed similar results for 16S rRNA gene libraries constructed from dry soils from Luther Vale. Their results also show complete absence of cyanobacteria from the low moisture soils. In a study by Aislabie et al. (2006), only one cyanobacterial ribotype was observed in a soil sample close to Lake Vanda in Victoria Land, which led to the assumption of wind dispersal. However, due to the high altitude of the 384 and 634 sample sites and their distance from lake Miers wind dispersal is very unlikely.



The cyanobacterial clones from libraries 171 and 184 identified showed overlap in terms of clusters, which is in contrast to the ARDRA groupings. Fewer groups are also present than in the cyanobacterial ARDRA profiling. The cyanobacterial

primer set could not amplify the clone A4 that was identified by BLAST as cyanobacterial. Only 33% of the Cyanobacteria classified could be resolved at family level with high confidence (≥80%) and all were assigned to Family 4.1.

A phylogenetic tree was constructed to analyse the distribution of the cyanobacterial sequences in samples 171 and 184 (Fig. 4.1). Groups were clustered from I to IV, excluding clones B17, A50 and B10 due to separate branching on the phylogenetic tree.

The B10 clone was the only sequence classified as *Oscillatoria*-like in this study. This genus is grouped under the order *Oscillatorialles* (subsection III) and is filamentous types present in both aquatic and terrestrial environments (Whitton and Potts, 2000). The clone B10 was completely separated from the rest of the clusters (Fig. 4.1). However, more *Oscillatoria*-like sequences must be identified to resolve their distribution in Dry Valley soils.



Cluster I was classified as *Cylindrospermum* with low confidence. Filamentous species from this genus have been found widespread in both aquatic and terrestrial environments (Whitton and Potts, 2000). This genus is grouped under the *Nostocales* order (subsection IV) and is very important in their natural habitat due to the capability to perform nitrogen fixation. Thirteen *Cylindrospermum*-like sequences were observed in this study. Eight clones were observed in cluster I with one (A13) and two clones (A47 and A8) each in mixed clusters II and IV respectively. The unclustered clones A50 and B17 were also classified as

*Cylindrospermum*-like. Six and seven clones each were observed in libraries 171 and 184 respectively corresponding to their widespread nature.

*Chroococcidiopsis*-like sequences (B79, B38, A15 and A82) were observed in mix clusters II and IV. Members of this genus are widespread cosmopolitan unicellular types that grow in aquatic and terrestrial environments (Fewer *et al.,* 2001). This genus is grouped under the *Pleurocapsales* order (subsection II) and is commonly found in extreme environments ranging from rocks from Antarctic Dry Valleys, hot desert soils, hypersaline environments and hot springs. These species can survive prolonged desiccation and are among the most ionizing-radiation-resistant bacteria known (Billi *et al.,* 2000). *Chroococcidiopsis* have been proposed as the most primitive living cyanobacterium (Friedmann *et al.,* 1994) but also as a candidate organism for terraforming Mars (Friedmann and Ocampo-Friedmann, 1995). In this study, 4 *Chroococcidiopsis*-like sequences were observed one clone per library (171 and 184). They formed part of mix clusters (II and IV) which grouped with *Cylindrospermum*-like clones on the phylogenetic tree (Fig. 4.1). However, Fewer et al. (2001) also observed the phenomenon of *Chroococcidiopsis* that formed close relationships with filamentous Cyanobacteria on a phylogenetic tree.

Cluster III was classified with low confidence as *Anabaena*, which is a freshwater genus sometimes present in brackish and marine waters but also present in near water sediments (Whitton and Potts, 2000). This genus is grouped under the

Cyanobacterial order *Nostocales* in subsection IV and posses the machinery to perform nitrogen fixation. The group consists of filamentous Cyanobacteria capable of forming branches. In this study, 6 *Anabaena*-like sp. were observed from samples 171 and 184. Only one of the clones (B13) was observed in library 184, which was 50 m away from the near lake Miers soil sample (171). This suggests that cluster III may harbour a lake specific group similar to *Anabaena*.

Acidobacteria followed by Actinobacteria dominated the other two samples 384 and 634. Acidobacteria is a newly recognized bacterial phylum found widespread in soil biotopes (Barnes *et al.,* 1999). They form a major part of non-cultured bacteria with only three cultured representatives: *Acidobacteria capsulatum*, *Halophaga foetida* and *Geothrix fermantans*. The majority of sequences identified from this division are from environmental clones. The phylum was originally divided into 8 subdivisions with subdivisions 1, 3, 4 and 6 represented only by environmental clone sequences (Hugenholtz *et al.,* 1998). However, Barnes et al. (2007) expanded the subdivisions to 26 in a recent study. The Acidobacterial dominance in the Miers Valley soils was also observe in the study by Smith et al. (2006). Interestingly, in this study the *gp4* and *gp6* identified clones were separated on the phylogenetic tree (Fig. 4.2). Barnes et al. (1999) constructed a phylogenetic tree using Acidobacterial environmental sequences and found that the *gp6* clones grouped with *Halophaga foetida* and *Geothrix fermantans* on the deepest branch of the tree. The *gp4* clones were found in all four libraries but *gp6* clones were specific to samples 171 and 184. A more detailed study on

Acidobacteria in Miers Valley soils may answer the question of group specificity in Antarctic soils.

Actinobacteria are heterotrophic high GC rich Gram-positive bacteria, which are well represented in culture studies (Basilio *et al.,* 2003). They are a well-defined phylogenetic group with only a small number of genera unknown. These poorly known genera are highly divergent from the other members of Actinobacteria and are classified as subclasses (Stackebrandt *et al.,* 1997). These subclasses are small, each represented only by a few strains, and include Rubrobacteridae, Acidimicrobidae, Sphaerobacteridae and Coriobacteridae (Holmes *et al.,* 2000). In this study Actinobacteria were only identified in samples 184, 384 and 634 with Rubrobacteridae (clones B61, D74 and D37) present in samples 184 and 634 and Pseudonorcadidae present in sample 384 (clones C20, C62 and C66). Interestingly, sequences classified in the group unclassified were also present only in samples 384 and 634 indicating a possible niche for novel species.

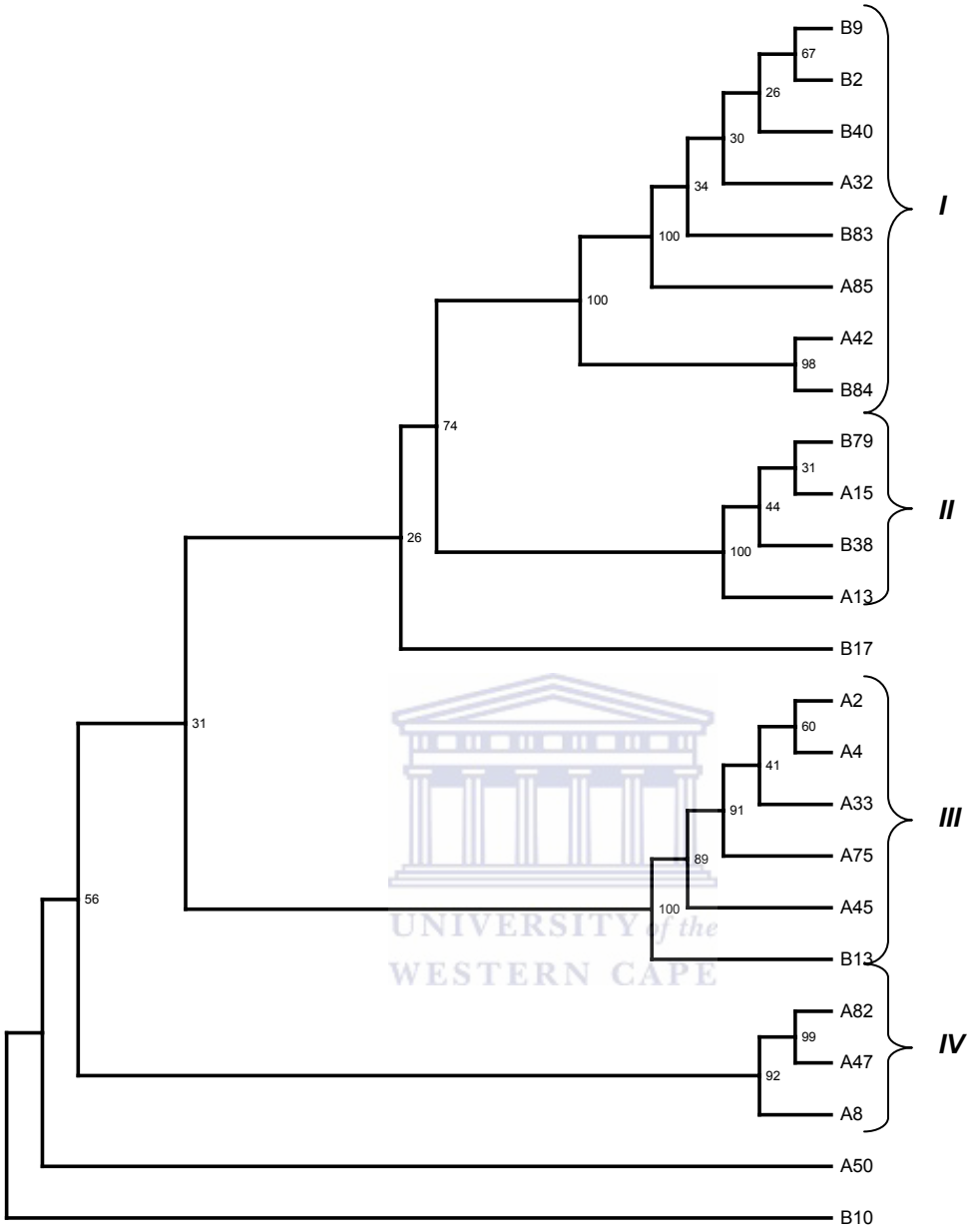


**Figure 4.1: Neighbour-joining tree constructed using cyanobacterial 16S rDNA sequences of 700 bp in length with the Juke and Cantor distances.** The numbers show the bootstrap values (higher than 50) from 1000 repetitions. I, II, III and IV denominate the cluster groups.

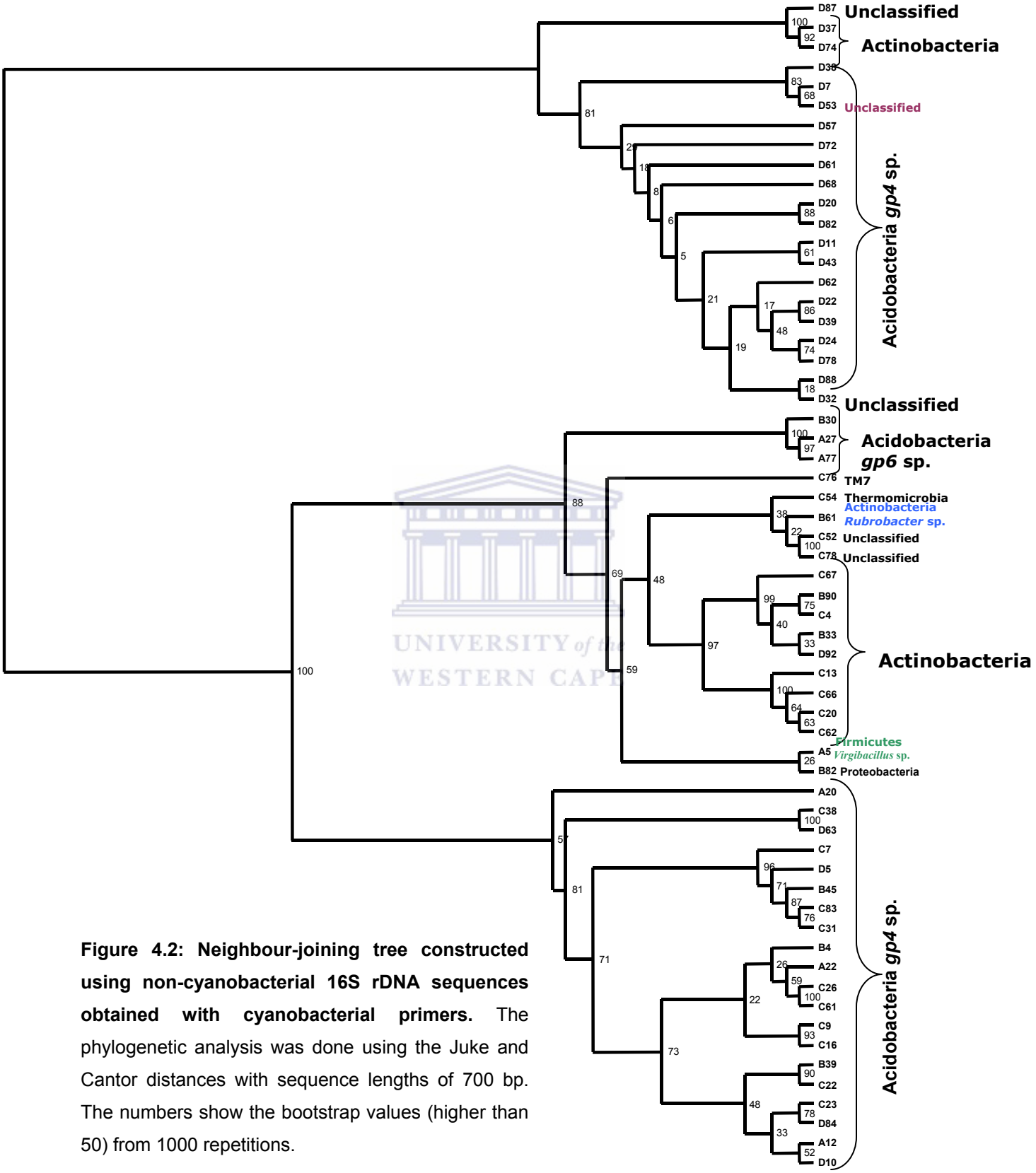


**Figure 4.2: Neighbour-joining tree constructed using non-cyanobacterial 16S rDNA sequences obtained with cyanobacterial primers.** The phylogenetic analysis was done using the Juke and Cantor distances with sequence lengths of 700 bp. The numbers show the bootstrap values (higher than 50) from 1000 repetitions.

# CHAPTER FIVE

## Conclusion and Future work

In conclusion, cyanobacterial distribution was found to be very interesting along the Miers Valley transect. The DGGE profiles of all eleven sites indicate presence of putative phylotypes whereas 16S rDNA libraries revealed cyanobacterial sequences present only in samples 171 and 184. The DGGE profiles indicated that some bands are specific to altitudinal sites. Other bands were observed in all samples. 16S rDNA Sequencing results indicate that cyanobacteria are only present in Antarctic lake and near lake sediments, which harbour moisture conditions appropriate for their growth. This result was similar to that observed by Niederberger et al. (2007). The sequencing also revealed a uniform distribution of some of the sequenced clones in the libraries (171 and 184). Parameters such as altitude, soil moisture, RH and soil texture could influence the distribution of Cyanobacteria in Antarctic soil biotopes.

Studies on Miers Valley soil by Smith et al. (2006) yielded high sequence amounts of the Acidobacterial phylum. This result was also observed in this study, which forces us to rethink the notion that Cyanobacteria dominate Antarctic soil biotopes. However, in this study Cyanobacteria specific primers were used which may have a more biased effect on the true composition of Acidobacteria in Antarctic soils.

The cyanobacterial specific primers amplified other genera, which affected subsequent experiments such as ARDRA analysis and richness estimations. This non-specific nature of the primers was proved using the RDP Probe Match tool, which revealed that mismatches could result in other bacterial genera being amplified.

Future investigations could involve the sequencing of more clones from the 171 and 184 16S rDNA libraries. We recommend the sequencing of the band types observed on the DGGE profile in order to confirm that the bands are cyanobacterial. The Nübel et al. (1997) primer set for cyanobacteria needs to be revised and redesigned. The optimization of PCR conditions should be carried out to achieve very specific amplification. Future work should also involve a detailed study on Acidobacteria from Antarctic soils, by both culturing and 16S rDNA analysis to understand the processes and cycles they are involved in.

# REFERENCES

Adams DG, Duggan PS (1999) Heterocyst and akinete differentiation in cyanobacteria. Tansley Review No. 107, New Phytol 144:3-33.

Aislabie JM, Chhour K, Saul DJ, Miyauchi S, Ayton J, Paetzold R, Balks MR (2006) Dominant bacteria in soils of Marble Point and Wright Valley, Victoria Land, Antarctica. Soil Biol Biochem 38:3041-3056.

Avrahami S, Liesack W, Conrad R (2003) Effects of temperature and fertilizer on activity and community structure of soil ammonia oxidizers. Environ Microbiol 5:691-705.



Baker JH, Smith DS (1972) The bacteria in an Antarctic peat. J Appl Bacteriol 35:589-596.

Bano N, Hollibaugh JT (2002) Phylogenetic composition of bacterioplankton assemblages from the Arctic Ocean. Appl Environ Microbiol 68:505-518.

Barns SM, Takala SL, Kuske CR (1999) Wide distribution and diversity of members of the bacterial kingdom Acidobacterium in the environment. Appl Environ Microbiol 65:1731-1737.

Barns SM, Cain EC, Sommerville L, Kuske CR (2007) Acidobacteria phylum sequences in Uranium-contaminated subsurface sediments greately expand the known diversity within the phylum. Appl Environ Microbiol 73:3113-3116.

Basilio A, Conzáles I, Vicente MF, Gorrochategui J, Cabello A, Gonzáles A, Genilloud O (2003) Patterns of antimicrobial activities from soil actinomycetes isolated under different conditions of pH and salinity. J Appl Microbiol 95:814-823.

Bermingham N, Luettich K (2003) Polymerase chain reaction and its applications. Curr Diagnos Pathol 9:159-164.

Billi D, Friedmann EI, Hofer KG, Grilli-Caiola M, Ocampo-Friedmann R (2000) Ionizing-radiation resistance in the dessication-tolerance cyanobacterium Chroococcidiopsis. Appl Environ Microbiol 66:1489-1492.

Boutte C, Grubisic S, Balthasart P, Wilmotte A (2005) Testing of primers for study of cyanobacterial molecular diversity based on DGGE. J Microbiol Methods 1-6.

Brambilla E, Hippe H, Hagelstein A, Tindall BJ, Stackebrandt E (2001) 16S rDNA diversity of cultured and uncultured prokaryotes of a mat sample from Lake Fryxell, McMurdo Dry Valleys, Antarctica. Extremophiles 5:23-33.

Brinkmann M, Pearce DA, Convey P, Ott S (2007) The cyanobacterial community of polygon soils at an inland Antarctic nunatak. Polar Biol 1-6.

Bull AT, Stach JEM (2004) An overview of biodiversity – Estimating the scale. In: Microbial diversity and bioprospecting, ed. Bull AT, ASM Press, Washington, DC. pp 15-28.

Cameron RE (1969) Cold desert characteristics and problems relevant to other arid lands. In: Arid lands in perspective, eds. McGinnies WG, Goldman BJ, Am Assoc Adv Sci pp 167-205.

Cameron RE, Devaney JR (1970) Antarctic soil algal crusts. A scanning electron and optical microscope study. Trans Am Microsc Soc 80:264-273.



Cameron RE (1971) Antarctic soil microbial and ecological investigations. In: Research in the Antarctic, eds. Quam LO, Porter HD, Am Assoc Adv Sci pp 137-189.

Cameron RE, Morelli FA, Randall LP (1972) Aerial aquaric and soil microbiology of Don Juan Pond, Antarctica. Antarctic J US 7:254-258.

Castenholz RW, Waterbury JB (1989) Chapter 2. Cyanobacteria in the environment. In: Bergey's Manual of Systematic Bacteriology Vol 3, eds. Staley JT, Bryant MP, Pfenning N, Holt JG, pp 1710-1727.

Cavacini P (2001) Soil algae from northern Victoria Land (Antarctica). Polar Bioscience 14:45-60.

Chao A (1984) Non-parametric estimation of the number of classes in a population. Scand J Statistics 11:265-270.

Christner BC, Kvitko BH, Reeve JN (2003) Molecular identification of bacteria and Eukarya inhabiting an Antarctic cryoconite hole. Extremophiles 7:177-183.



Clark MS, Clarke A, Cockell CS, Convey P, Detrich HW, Fraser KPP, Johnston IA, Methe BA, Murray AE, Peck LS, Römisch K, Rogers AD (2004) Antarctic genomics. Comp Funct Genom 5:231.

Cowan DA, Russell NJ, Mamais A, Sheppard DM (2002) Antarctic Dry Valley mineral soils contain unexpectedly high levels of microbial biomass. Extremophiles 6:431-436.

Cowan DA, Ah Tow L (2004) Endangered Antarctic environments. Annu Rev Microbiol 58:649-690.

Crump BC, Kling GW, Bagar M, Hobbie JE (2003) Bacterioplankton community shifts in an arctic lake correlate with seasonal changes in organic matter source. Appl Environ Microbiol 69:2253-2286.

de la Torre JR, Goebel BM, Friedmann EI, Pace NR (2003) Microbial diversity of cryptoendolythic communities from the McMurdo Dry Valleys, Antarctica. Appl Environ Microbiol 69:3858-3867.

de los Ríos A, Ascaso C, Wierzchos J, Fernández-Valiente E, Quesada A (2004) Microstructural characterization of cyanobacterial mats from the McMurdo ice shelf, Antarctica. Appl Environ Microbiol 70:569-580.

Desai U, Phaffle P (1995) Single-step purification of thermostable DNA polymerase expressed in *Escherichia coli*. Biotechniques 19:780-784.

Doran PT, Priscu JC, Lyons WB, Walsh JE, Fountain AG, McKnight DM, Moorhead DL, Virginia RA, Wall DH, Clow GD, Fritsen CH, McKay CP, Parsons AN (2002) Antarctic climate cooling and terrestrial ecosystem response. Nature 415:517-520.

Dunbar J, Barns SM, Ticknor LO Kuske CR (2002) Empirical and Theoretical Bacterial Diversity in four Arizona soils. Appl Environ Microbiol 68:3035-3045.

Ercolini D (2004) PCR-DGGE fingerprinting: novel strategies for detection of microbes in food. J Microbiol Methods 56:297-314.

Fewer D, Friedl T, Büdel B (2002) *Chroococcidiopsis* and heterocyst-differentiating Cyanobacteria are each other's closing living relatives. 23:82-90.

Fierer N, Jackson RB (2006) The diversity and biogeography of soil bacterial communities. PNAS 103:626-631.

Friedmann EI (1993) Antarctic microbiology. Wiley-Liss, Inc., New York. pp 1-615.

Friedmann EI, Ocampo-Friedmann R, Hua M (1994) *Chroococcidiopsis*, the most primitive living cyanobacterium? Origins Life Evol, Biosphere 24:269-270.

Friedmann EI, Ocampo-Friedmann R (1995) A primitive cyanobacterium as pioneer microorganism for terraforming Mars. Adv Space Res 15:243-246.

Fritsen CH, Adams EE, Mckay CM, Priscu (1998) Permanent ice covers of the McMurdo Dry Valley Lakes, Antarctic liquid water content. In: Ecosystem Dynamics in a Polar desert: The McMurdo Dry Valleys, Antarctica, ed. Priscu JC. Antarctic Research Series 72:269-280.

Gallon JR, Jones DA, Page TS (1996) *Trichodesmium*, the paradoxical diazotroph. Algol Stud 83:215-243.

Gans J, Wolinsky M, Dunbar J (2005) Computational improvements reveal great bacterial diversity and high metal toxicity in soil. Science 309:1387-1390.

Gentry TJ, Wickham GS, Schadt CW, He Z, Zhou J (2006) Microarray applications in microbial ecology research. Microbial Ecol 52:159-175.

George AL, Murray AW, Montiel PO (2001) Tolerance of Antarctic cyanobacterial mats to enhanced UV radiation. FEMS Microbiol Ecol 37:91.

Gordon DA, Priscu JC, Giovannoni S (2000) Distribution and phylogeny of bacterial communities associated with mineral particles in Antarctic lake ice. Microbial Ecol 39:197-202.

Hansen M, Nielson-Tolker T, Givskov M, Molin S (1998) Biased 16S rDNA PCR amplification caused by interference from DNA flanking region. FEMS Microbiol Ecol 26:141-149.

Hayes VM, Wu Y, Osinga J, Mulder IM, van der Vlies P, Elfferich P, Buys CHCM, Hofstra RMW (1999) Improvements in gel composition and electrophoretic

conditions for broad-range mutation analysis by denaturing gradient gel electrophoresis. Nucleic Acids Res 27:1-4.

Holmes AJ, Bowyer J, Holley MP, O'Donoghue M, Montgomery M, Gillings MR (2000) Diverse, yet-to-be-cultured members of the *Rubrobacter* subdivision of the Actinobacteria are widespread in Australian arid soils. FEMS Microbiol Ecol 33:111-120.

Hong H, Pruden A, Reardon KF (2007) Comparison of CE-SSCP and DGGE for monitoring a complex microbial community remediating mine drainage. J Micorbiol Methods 69:52-54.

Horowitz NH, Cameron RE, Hubbard, JS (1972) Microbiology of the Dry Valleys of Antarctica. Antarct Sci 176:242-245.

Hugenholtz P, Goebel BM, Pace NR, (1998) Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. J Bacteriol 180:4765-4774.

Jackson CR, Roden EE, Churchill PF (2000) Denaturing gradient gel electrophoresis can fail to separate 16S rDNA fragments with multiple base differences. Mol Biol Today 1:49-51.

Janssen PH, Yates PS, Grinton BE, Taylor PM, Sait M (2002) Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions *Acidobacteri*a, *Actinobacteri*a, *Proteobacteri*a, and *Verrucomicrobia*. Appl Environ Microbiol 68:2391-2396.

Kappen L, Straka H (1988) Pollen and spores transport into the Antarctic. Polar Biol 8:173-180.

Kuske CR, Barns SM, Busch JD (1997) Diverse uncultivated bacterial groups from soils of the arid southwestern United States that are present in many geographic regions. Appl Environ Microbiol 63:3614-3621.

Letowski J, Brousseau R, Masson L (2004) Designing better probes: effect of probe size, mismatch position and number on hybridization in DNA oligonucleotide microarrays. J Microbiol Methods 57:269-78.

Liesack W, Stackebrandt E (1992) Occurrence of novel groups of the domain bacteria as revealed by analysis of genetic material isolated from an Australian terrestrial environment. J Bacteriol 174:5072-5078.

Line MA (1988) Microbial flora of some soils of Mawson Base and the Vestfold Hills, Antarctica. Polar Biol 8:421-427.

Madigan MT, Martinko JM, Parker J (2000) Brock Biology of Microorganisms. 9th edn. Prentice Hall International, New Jersey.

McDougald D, Rice SA, Weichart D, Kjelleberg S (1998) Nonculturability: adaptation or debilitation? FEMS Microbiol Ecol 25:1-9.

Miller DN, Bryant JE, Madsen EL, Ghiorse WC (1999) Evaluation and Optimization of DNA Extraction and Purification Procedures for Soil and Sediment Samples. Appl Environ Microbiol 65:4715-4724.

Miller SR, Bebout BM (2004) Variation in sulfide tolerance of photosystem II in phylogenetically diverse cyanobacteria from sulfidic habitats. Appl Environ Microbiol 70:736-744.



Muralidharan K, Wemmer C (1994) Transporting and storing field-collected specimens for DNA without refrigeration for subsequent DNA extraction and analysis. Biotechniques 17:420-422.

Muyzer G, de Waal EC, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol 59: 695-700.

Nadeau T, Howard-Williams C, Castenholtz RW (1999) Effects of solar UV and visible irradiance on photosynthesis and vertical migration of *Oscillatoria* sp. (Cyanobacteria) in an Antarctic microbial mat. Aquatic Microbial Ecol 20:231-243.

Nakagawa T, Nagaoka T, Taniguchi S, Miyaji T, Tomizuka N (2004) Isolation and characterization of psychrophilic yeasts producing cold-adapted pectinolytic enzymes. Letters in Appl Microbiol 38:383-389.

Neefs J, van de Peer Y, Hendriks L, de Wachter R (1990) Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Res 18:1-5.

Nemergut DR, Martin AP, Schmidt SK (2004) Integron diversity in heavymetal contaminated mine tailings and inferences about integron evolution[J], Appl Environ Microbiol 70:1160–1168.

Niederberger TD, McDonald IR, Hacker AL, Soo RM, Barrett JE, Wall DH, Cary CS (2007) Microbial community structure in soils of Northern Victoria Land, Antarctica. Environ Microbiol. Submitted.

Nielsen HB, Wernersson R, Knudsen S (2003) Design of oligonucleotides for micorarrays and perspectives for design of multi-transcriptome arrays. Nucleic Acids Res 31:3491-3496.

Nübel U, Garcia-Pichel F, Muyzer G (1997) PCR primers to amplify 16S rRNA genes from Cyanobacteria. Appl Environ Microbiol 63:3327-3332.

Orita M, Suzuki Y, Sekiya T, Hayashi K (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics 5:874-879.

Oto M, Suda W, Shinoyama H (2006) 16S rRNA gene-based analysis of microbial community by whole-genome amplification and minigel-single-strand conformation polymorphism technique. J Bioscience Bioengineering 102:482-484.



Pace NR, Stahl DA, Lane DJ, Olsen GJ (1986) The analysis of natural microbial populations by ribosomal RNA sequences. Adv Microb Ecol 9:1-55.

Priscu JC, Fritsen CH, Adams EE, Giovannoni SJ, Paerl HW, McKay CP, Doran PT, Gordon DA, Lanoil BD, Pinckney JL, (1998) Perenial Antarctic lake ice: an oasis for life in a polar desert. Science 280:2095-2098.

Reed RH, Chudek JA, Foster R, Stewart WDP (1984) Osmotic adjustment in cyanobacteria. Arch Microbiol 138:333-337.

Reysenbach A-L, Pace NR (1995) Reliable amplification of hyperthermophilic archaeal 16S rRNA genes by the Polymerase Chain Reaction. In: *Archaea*: A Laboratory Manual-Thermophiles, eds. Robb FT, Place AR. Cold Spring Harbour Laboratory Press, pp 101-107.

Rudi K, Skulberg M, Skulberg R, Jakobsen KS (2000) Application of sequence-specific labeled 16S rRNA gene oligonucleotide probes for genetic profiling of cyanobacterial abundance and diversity by array hybridization. Appl Environ Microbiol 66:4004-4011.

Russell NJ, Hamamoto T (1998) Psychrophiles. In: Extremophiles: Microbial life in extreme environments, eds. Horikoshi K, Grant WD. Wiley-Liss and Son, pp 25-45.



Sambrook J, Russell DW (2001) Molecular cloning: A laboratory manual. Cold spring harbour laboratory press, cold spring harbour, New York.

Schopf JW (1996) Cyanobacteria. Pioneers of the early Earth. In: Contributions in Phycology, eds. Prasad AKSK, Nienow JA, Rao NR. 13-32.

Sekiguchi H, Watanabe M, Nakahara T, Xu B, Uchiyama H (2002) Succession of bacterial community structure along the Changjiang river determined by

denaturing gradient gel electrophoresis and clone library analysis. Appl Environ Microbiol 68:5142-5150.

Simmons GW, Vestal JR, Wharton RA (1993) Environmental regulators of microbial activity in continental Antarctic lakes. In: Physical and Biogeochemical Processes in Antarctic Lakes, eds. Green WJ, Friedmann EI. Antarctic Res Series 59:197-214.

Smith HG, Tearle PV (1985) Aspects of microbial and protozoan abundances in Signy Island fellfields. Brit Ant Surv Bull 66:19-33.

Smith JJ, Ah Tow LA, Stafford W, Cary C, Cowan DA (2006) Bacterial diversity in three different Antarctic cold desert mineral soils. Microbial Ecol 51:413-421.



Stackebrandt E, Rainey FA, Ward-Rainey NL (1997) Proposal for a new hierarchic classification system, *Actinobacteria* classis nov. Int J Syst Bacteriol 47:479-491.

Stanier G (1988) Fine structure of cyanobacteria. In: Methods on enzymology, eds. Packer L, Glazer AN. Volume 167:157-172.

Steven B, Léveillé R, Pollard WH (2006) Microbial ecology and biodiversity in permafrost. Extremophiles 10:259-267.

Tang EPY, Tremblay R, Vincent WF (1997) Cyanobacterial dominance of polar freshwater ecosystems: Are high latitude-mat-formers adapted to low temperature? J Phycol 33:171-181.

Taton A, Grubisic S, Brambilla E, De Wit R, Wilmotte A (2003) Cyanobacterial diversity in natural and artificial microbial mats of Lake Fryxell (McMurdo Dry Valleys, Antarctica): a morphological and molecular approach. Appl Environ Microbiol 69:5157-5169.

Taton A, Grubisic S, Bathasart P, Hodgson DA, Laybourn-Parry J, Wilmotte A (2006) Biogeographical distribution and ecological ranges of benthic cyanobacteria in East Antarctic lakes. FEMS Microbiol Ecol 57:272-289.



Taylor G (1916) With Scott: The Silver Lining. Dodd, Mead and Co, New York.

Tindall BJ (2004) Prokaryotic diversity in the Antarctic: The tip of the iceberg. Microbial Ecol 47:271-283.

Tomitani A, Knoll AH, Cavanaugh CM, Ohno T (2006) The evolutionary diversification of cyanobacteria: Molecular-phylogenetic and paleontological perspectives. PNAS 103:5442-5447.

Torsvik V, Goksøyr J, Daae FL (1990) High diversity in DNA of soil bacteria. Appl Environ Microbiol 56:782–787.

Tsai YL, Olsen BH (1991) Rapid method for direct extraction of DNA from soil and sediments. Appl Environ Microbiol 57:1070-1074.

Urakawa H, Fantroussi EI, Smidt H, Smoot JC, Tribou EH, Kelly JJ, Noble PA, Stahl DA (2003) Optimization of single-base-pair mismatch determination in oligonucleotide micorarrays. Appl Environ Microbiol 69:2848-2856.

Vaneechoutte M, De Beenhouwer G, Claeys G, Verschraegen G, De Rouck A, Paepe N, Elaichouni A, Portaels F (1993) Identification of *Mycobacterium* species by using amplified ribosomal DNA restriction analysis. J Clin Microbiol 31:2061-2065.

Vincent WF, Howard-Williams C, Broady PA (1993) Microbial communities and processes in Antarctic flowing waters. In: Antarctic Microbiology, ed. Friedmann EI Wiley-Liss Inc., New York, pp 543-569.

Vishniac HS (1993) The microbiology of Antarctic soils. In: Antarctic Microbiology, ed Friedmann EI, New York, Wiley-Liss Inc., pp 297-341.

von Wintzingerode F, Göbel UV, Stackebrandt E (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS Microbiol 21:213-229.

Wing KT, Priscu JC (1993) Microbial communities in the permanent ice cap of Lake Booney, Antarctica: Relationship among Chlorophyl *a*, gravel and nutrients. Antarctic J, Review Volume 28:246-249.

Whitman WB, Coleman DC, Wiebe WJ (1998) Prokaryotes: The unseen majority. Proc Natl Acad Sci USA 95:6578-6583.

Whitton BA, Potts M (2000) The ecology of cyanobacteria, their diversity in time and space. 1st edn. Kluwer Academic Publishers, Dordrecht.



Wynn-Williams DD (1990) Ecological aspects of Antarctic microbiology. In: Advances in Microbial Ecology vol. 2, ed Marshall KC, Plenum Publishing, pp 71-146.

Xu J (2006) Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. Mol Ecol 15:1713-1731.

Zhou J, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. Appl Environ Microbiol 62:316-322.

Zwart G, Kampst-van Agterveld MP, van der Werff-Staverman I, Hagen F, Hoogveld HL, Gons HJ (2005) Molecular characterization of cyanobacterial diversity in a shallow eutrophic lake. Environ Microbiol 7:365-377.