METAGENOMICS AND COMPOSITION OF PROKARYOTES ON LEWIS GLACIER AND ITS FORELAND IN MOUNT KENYA, NYERI COUNTY

JOSIAH OCHIENG KUJA

DOCTOR OF PHILOSOPHY

(Microbiology)

JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY

2019

Metagenomics and Composition of Prokaryotes on Lewis Glacier and its Foreland in Mount Kenya, Nyeri County

Josiah Ochieng Kuja

A Thesis Submitted in Fulfillment for the Degree of Doctor of Philosophy in Microbiology in the Jomo Kenyatta University of Agriculture and Technology

2019

DECLARATION

This thesis is my original work and has not been presented for a degree in any other University.

Signature ..... Date.....

Josiah Ochieng Kuja

This thesis has been submitted for examination with our approval as university supervisors.

Signature ..... Date.....

Prof. Hamadi Iddi Boga, PhD

JKUAT, Kenya

Signature ..... Date.....

Prof. Anne W. T. Muigai, PhD

JKUAT, Kenya

Signature ..... Date.....

Dr. Huxley Mae Makonde, PhD

Technical University of Mombasa, Kenya

Signature ..... Date.....

Dr. Jun Uetake, PhD

Colorado State University, USA

DEDICATION

This work is dedicated to my dear family; my wife Alice Kuja, daughter; Jun Kuja, my mother; Lucy Kuja, late father Peter Kuja, my brothers and sisters, baba Kinya Munyirwa, aunt Tabitha Kinya and the entire Munyirwa family. I appreciate the support you have accorded me during the course of my studies. Without your encouragement and support this journey would have been longer and tougher.

ACKNOWLEDGEMENT

I thank Almighty God for binging me this far from the commencement of this long and challenging journey.

I wish to thank the following institutions for their assistance in this research. The Japanese Society for the Promotion of Science (JSPS) and the National Commission for Science, Technology and Innovation (NACOSTI), for a research grant and permit; Kenya Wildlife Service (KWS) for access permit to the sampling area in Mt. Kenya; and National Institute of Polar Research (NIPR) for the lab space and equipment.

I wish to acknowledge my supervisors, Prof. Hamadi Boga for the introduction into the scientific leadership, Prof. Anne Muigai for her cordial support and spiritual guidance, Dr. Huxley Mae Makonde for his unavoidable reach and inspiration and Prof. Jun Uetake for the introduction to the cryophilic and bioinformatics world. I also wish to thank NIPR technical support, Kenichio Wetanabe, and Dr. Amanuel for their support and guidance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND ACRONYMS

ANOSIM	Analysis of Similarity
CDD	Conserved Domain Database
DNA	Deoxyribonucleic Acid
GPS	Global Positioning System
KEGG	Kyoto Encyclopedia of Genes and Genomes
MEGAN	Metagenome Analyzer
NAST	Nearest Alignment Space Termination
NCBI	National Centre for Biotechnology Information
NMDS	Non Metric Dimensional Scaling
OTU's	Operational Taxonomic Units
ASV	Amplicon Sequence Variance
PCR	Polymerase Chain Reaction
R2A	Reasoner's 2A medium
BG	Blue-Green medium
PyNAST	Python Nearest Alignment Space Termination
QIIME	Quantitative Insights into Microbial Ecology
RDA	Redundancy analysis
RNA	Ribonucleic Acid
SDS	Sodium Dodecyl Sulphate
SSU	Small Sub Unit
UCLUST	Universal Clustering
ARISA	Automated Ribosomal Space Analysis
BLAST	Basic Local Alignment Search Tool
EC	Electrical Conductivity
GPS	Global Positioning System
HTS	High Throughput Sequencing
NCBI	National Centre for Biotechnology Information
NMDS	Non-metric multidimensional Scaling

ppm	Parts per million
QIIME	Qualitative Insight into Microbial Ecology
CAP	Correspondence Analyses of Proximities
PCoA	Principal Component Analyses
RDA	Redundancy Analysis
RDP	Ribosomal Database Project
SSU	Small Subunit
T (°C)	Temperature in degree Celsius
TRFLP	Terminal Restriction Fragment Length Polymorphism

ABSTRACT

The Lewis glacier in Mt. Kenya, Kenya is one of the best documented tropical glaciers globally, with about 80 years of frequent observations of the length, area and volume change. These studies have, however, focused on the measurements of the surface mass balance and the recession of Lewis Glacier. Currently the Lewis Glacier is less than 0.105 ± 0.001 × 106 m², indicating a massive loss of the glacier mass and its endogenous biodiversity. The main objective of this study was to isolate and characterize glacial prokaryotic populations and determine their community structures on Lewis glacier using the next generation sequencing techniques by targeting the 16S rDNA V3-V4 variable gene regions. Samples collected included cryoconite materials, foreland soil and foreland plant rhizosphere from 12-year-old, 23-year-old and 42-year-old foreland chronosequence. R2A and BG-11 media were used to inoculate melt water into a hundred folds. Inoculants were incubated at 4°C, 10°C, and 25°C. Total community DNA was extracted from samples using phenol-chloroform. The 16S rDNA gene variable region (V3 – V4) of the extracted DNA library construction was performed as per the Illumina sequencing protocol. Sequences were analysed using QIIME2 pipeline. The isolates were partially sequenced by Sanger and the sequences aligned to their closest relatives from the NCBI gene bank. The QIIME2 output were subsequently introduced into R programming language (3.5.1) for the analyses of species abundance, composition, diversity and associations between the bacterial community structure and the environmental constraints. Packages tidyverse, superheat, vegan, dplyr, indicspecies, indval were used for diversity, composition, ordinations indicator species analyses. Packages ggrepel, ggcorrplot, were used for the multivariate analyses. The concentration of dissolved organic matter and nutrients were measured as eluted supernatant of the 10% UWA medium using a Spectroquant test kit and a spectral photometer. All samples were filtered through a polycarbonate membrane. The nutrient eluted supernatant of the 10% UWA agar medium were collected after shaking (110 r.p.m.) for 24 h. Albedo reflectance was measured in five replicates using albedo-meter from the field sites. Phylogenetic analyses of the 16S rDNA of axenic isolates revealed three major phyla *Firmicutes, Proteobacteria,* and *Actinobacteria* to dominate the snowpack. *Cyanobacteria* was the most prominent phylum (35 % - 37 %) on the glacier. The distribution of these genera were significantly different between the upper and the lower glacier ($r = 0.094$, $p = 0.027$) and between the foreland soil and rhizosphere ($r = 0.27$, $p = 0.001$). Other phyla, *Proteobacteria, Actinobacteria* and *Bacteroidetes* were also prominent in the glacier and at the foreland chrono-sequence. The community structure and distribution of these phyla were significantly different in the sample types ($r = 0.84$, $p = 0.001$), age of the last glacier terminal ($r = 0.707$, $p = 0.001$), sample sites ($r = 0.816$, $p = 0.001$) and altitude ($r = 0.63$, $p = 0.011$). There were a total of 27,367 ASVs across the glacier and foreland ecosystem. 185 ASVs were shared across the glacier, rhizosphere and soil samples, 423 ASVs were shared across the 12-year-old, 23-year-old and 42-year-old foreland rhizosphere while 455 ASVs were shared across the 12-year-old, 23-year-old and 42-year-old foreland soil. Foreland plant rhizosphere had the highest microbial community structure that increased down the foreland age. Primary

foreland had the highest number of ASVs while glacier sites had the lowest microbial community structures, but high ASV numbers. The 3 % distance analyses of the of the 16S rDNA revealed that the rarefaction sequencing depth is far from exhaustive sampling in a number of samples, including the largest samples. Ordination analyses showed samples to be clustered together according to the sample types and sampling sites. This was also supported by the correspondence analyses of proximities (CAP) between the environment and bacterial community matrices. The analyses of correlation between the environmental variables and bacterial community structure indicated cryoconite material and mineral to be strongly correlated ($r = 1$). The Euclidean distance and Bray-Curtis dissimilarity were maximally correlated and revealed altitude and Chlorophyll.a to strongly influence the bacterial community structure ($r = 0.318$). Mantel test further confirmed the relationship between the environmental factors to bacterial community structure ($r = 0.1763, p = 0.004$). Like any other glacier in the world, the Lewis Glacier is dominated with the phylum *Cyanobacteria* which is less distributed across the foreland chrono-sequence. Other phyla such as *Proteobacteria, Actinobacteria* and *Bacteroidetes* also inhabits the glacier ecosystem, but highly abundant in the foreland. The current results are important as reference database tropical cryophilic prokaryotes. The study has also led to the conservation of low temperature adapted microorganisms in the local laboratory for subsequent analyses and screening for biotechnological applications.

CHAPTER ONE

INTRODUCTION

1.1.General Background to the study

A Glacier is a persistent body of ice which flows downwards by their weight (Bell *et al.,* 2011). Ninety-nine percent of glacier ice is located in the bipolar region and remains are in high altitude regions (Moreno *et al.,* 2001) including the tropical Africa continent. Until recently, glaciers have long been believed to be almost abiotic environments. However, glaciers are relatively fertile ecosystems with many psychrophilic organisms such as snow algae (Hoham and Duval*,* 2001; Kol, 1969), cyanobacteria (Takeuchi *et al.,* 2000), yeast (Branda et al., 2010a), bacteria (Segawa *et al.,* 2005), invertebrates and metazoans (Christner *et al.,* 2003).

The occurrence of these organisms is characterized with the cryoconite granules that contain organic and inorganic particles (Takeuchi *et al.,* 2001; Takeuchi *et al.,* 2010). Organic particles result from the active biological activities that take place on glacial niche which might lead to glacial melting (Takeuchi *et al.,* 2001) and disappearance (Prinz *et al.,* 2011a). The organic and inorganic substances (Takeuchi *et al.,* 2001) significantly contribute to glacier surface energy heat budget, glacier avalanche and channelled melt water runoff redistribution (Kustas *et al.,* 1994; Paterson, 1994; Hock & Holmgren, 2005) to the glacier bed and pressure melting points (Hodson *et al.,* 2008a).

The population and functional roles of the microbial ecology in glacial zones are regulated by the physio-chemical activities in microbial chains involving nutrient utilization (Sigler *et al.,* 2002). Spatial biogeochemical processes within the cryoconite material also contribute to microbial functions and adaptability (Stibal *et al.,* 2012; Telling *et al.,* 2012). The activity of these microbial chains is also possible due to the available melt water, debris and aerosol deposits (Hodson *et al.,* 2008b; Stibal *et al.,* 2012) as a result of endogenous and exogenous processes including episodic atmospheric processes (Kühnel

et al., 2011; Kekonen *et al.,* 2005), seasonal external environment (Hodson *et al.,* 2008a) and the microbial activities on the glacier surfaces (Langford *et al.,* 2010; Stibal *et al.,* 2012).

In Africa there are three higher mountains to the East of Africa where glacier exist. They include Mount Kilimanjaro (5899m; Tanzania), Mount Kenya (5199m; Kenya) and Mount Rwenzori (5109m; Uganda). The existing glacier zones vary in size due to rapid melting (Thomson *et al.,* 2009; Rabatel *et al.,* 2013). The recession of the East African glacier is dated to early 1880s (Hastenrath, 1983a; Hastenrath, 2006). However, the presence of glacier in the African continent is evidenced by the presence of Pleistocene areas (Anwar, Flickr, & Kilimanjaro, n.d.). The high mountains are characterised by the structured volcanoes.

Mount Kilimanjaro in Tanzania is the highest free standing mountain in the world (Duane *et al.,* 2008) characterised with three volcanoes: Kibo, Mawenzi, and Shira (Downie, 1964). It had fifteen glaciers and three ice caps that surrounded Kibo peak (Mote & Kaser, 2007). These glaciers and the ice caps are estimated to have reduced from 20 Km^2 in 1800 to 2.51 Km^2 in 2003 (Cullen *et al.,* 2006). The dramatic lose has been linked to climate change and rapid human activities including deforestation along the mountain contours (Mölg *et al.,* 2003).

Mount Rwenzori, in Uganda, is the third highest mountain in Africa. It is characterised with three peaks: Stanley, Speke and Baker, that still contain glaciers (Klein & Kincaid 2007). Stanley is the highest peak, 5000 m (Young & Hastenrath, 1991). In 2014, a large black bio- aggregating matter was identified by Uetake et al. (2014), as a glacier moss gemmae aggregation (GMGA) which occurred on the Rwenzori glacier surface. The biodiversity on the Rwenzori glaciers are threatened due to the rapid shrinking, which was evidenced by Young & Hastenrath (1991). These glaciers have, however, disappeared by melting due to irregular air temperatures, less snow accumulation and vaporization attributed to decreased cloud cover (UNEP, 2005).

Mount Kenya, located in Kenya, East Africa, is the highest mountain in Kenya and the second highest mountain in Africa (5199 m) after mount Kilimanjaro (5899 m). Mount Kenya was covered with 18 glaciers: Darwin, NW Pigott, Barlow, Cesar, Josef, Peter, Northey, Mackinder, Arthur, Diamond, Foreil, Heim, Tyndal, Melhuish, Krapf, Kolbe, Gregory and Lewis glacier (Young & Hastenrath, 1991; Hastenrath, 2008), which gradually decreased by shrinking (Prinz *et al.,* 2011a) and rapidly from 1934 (Thomson *et al.,* 2009; Rabatel *et al.,* 2013). An inventory in 2004 showed that, the 18 glaciers had reduced from 1.64 km^2 at the end of 19^{th} century to 0.27 km^2 in 2004 (Prinz *et al.,* 2018). Eight glacier had, however, completely disappeared (Hastenrath *et al.,* 2005) in the 20^{th} century by suffering a substantial loss (Prinz *et al.,* 2012a) and only a few, Lewis and Tyndal glacier are surviving. Even Gregory glacier that was joined to Lewis glacier from the leeward side was the last to suffer rapid loss in 2011 (Prinz et al., 2011a).

Lewis glacier is the largest glacier in mount Kenya that is still surviving and well-studied. Most of the studies have, however, inclined to real time remote sensing of the glacier recession based on time as age (Prinz *et al.,* 2018). The high resolution imaging of Lewis glacier has revealed a tremendously reduction (90%) of its volume between 1934 and 2010 (Prinz *et al.,* 2010, 2011a, 2012, 2018). Recent climate sensitivity study by Prinz et al. (2016) suggested that the glacier recession may be due to atmospheric drying resulting from less cloud cover, reduced snowfalls and reduced albedo. This suggestion concurred with studies elsewhere (Vuille *et al.,* 2008; Mölg *et al.,* 2009) where dry atmosphere and low cloud cover are linked to constant reduction of the size and shape of glaciers. Lewis glacier is also characterised by the cryoconite holes and surface dust particles. It also experiences the Leeward dust particle influxes and windward predominant weather from the Indian ocean (Kuja *et al.,* 2018a).

Lewis glacier is structured into two sections at 4819 m above sea level and 4670 m above sea level with slope gradients of 17.4 degrees and 19.8 degrees respectively (Kuja *et al.,* 2018). The cryoconite material in the Lewis glacier might be harbouring complex extracellular polymeric substances and variable microbial community like other glaciers

(Hodson *et al.,* 2008b) making it act as a micro-biogeochemical reactor (Hodson *et al.,* 2008a; Stibal *et al.,* 2012) that is highly exposed to deglaciation. The biology of Lewis glacier is, however, poorly understood with no attempts to prior investigations on the possible microbial ecology and biodiversity. In addition to the geological inventories (Young & Hastenrath, 1991; Hastenrath, 2008; Prinz *et al.,* 2010, 2011, 2012, 2018), there has been a progressive studies on the foreland succession studies on the low and high vascular plant cover (Mizuno & Fujita, 2014). The rapid loss of Lewis glacier (Rabatel *et al.,* 2013) threatens the possible biodiversity, especially microbial community structures that might be of commercial and industrial significance. The analyses of the abundance, diversity and the composition of prokaryotic communities from Lewis glacier is important to conservation of threatened novel organisms.

1.2.Statement of Problem

Glaciers are biotic with most studies focussing on the polar, alpine and mid-latitude regions. They harbour unique microbial communities with significant contribution to cryoconite aggregation through the production of cohesive extracellular polymeric substances (Sigler & Zeyer, 2002b). The glacier biology of the tropical African glaciers have not been investigated to establish the abundance, diversity and composition of the microbial community structures.

The surviving tropical African glacier are rapidly disappearing by melting. The rapid loss of the glacier content is a threat to the glacier endogenous biodiversity. Moreover, the surviving African glaciers lie astride the equator exposing the glacier surface to constant sunlight radiations which are absorbed into glacier content. This effect is relatively low in the polar glacier, which are characterised with the accumulation layer of the glacier.

The Lewis Glacier on Mt Kenya is slightly less than 0.4 km^2 tropical glacier with no accumulation supra-glacial zones (Prinz *et al.,* 2012b). The lack of the accumulation layer significantly contributes to the glacial shrinking and disappearance. The phenomenon subsequently contributes to rapid loss of the psychrophilic and psychotropic endogenous

biodiversity from the tropical glacier ecosystem through channelized surface water tillage. In addition, Lewis glacier is structured into two sections at 4819 m above sea level and 4670 m above sea level with slope gradients of 17.4 degrees and 19.8 degrees respectively. The differences in the slope of the glacier sites facilitates the rate of glacier avalanche and channelized melt water runoff redistribution. These processes also affect the residence time within the cryoconite holes.

The cryoconite holes contain cryoconite materials and accumulated surface dust particles forming dark particles on the glacier surface. The dark glacier surface is a recipe to reduced radiation reflectance on the glacier surface. The reduced surface reflectance increases the supra-glacier surface heat budget. This makes cryoconite holes act as a micro-biogeochemical reactor that is highly exposed to rapid melt water tillage.

All these factors contribute to the stability and residence time of the cryoconite holes, that contain the microbial communities. The biodiversity in Lewis glacier is therefore threatened by the rapid disappearance of the glacier content in to the foreland ecosystem. The present study therefore aimed at the isolation, characterization and conservation of the threatened prokaryotes from the glacier surface through determination of the taxa diversity, abundance and composition within the cryoconite samples. The foreland ecosystem was also studied to establish the dynamics between the glacier and foreland microbial community structures. This would lead to the establishment of the variation in microbial diversity, their community structure, composition between the glacier and foreland chrono-sequence.

1.3.Justification of the study

In Africa biodiversity and ecological studies have focused on the low altitude terrestrial and aquatic ecosystems. The biological studies in these regions are well established with reference to novel microorganism in molecular biology and bioengineering. The high altitude tropical glacier ecosystems have remained neglected for decades regardless of their rapid loss. Usually, the high mountains containing the glacier have been believed to

be abiotic. Similar regions like the alpine and polar glaciers are known to harbour diverse microorganisms such as bacteria, fungi, archaea, and viruses. A number of biological studies from these regions (Hoham & Duval, 2001; Takeuchi et al., 2001; Turchetti *et al.,* 2008), Segawa *et al.,* 2005; Christner *et al.,* 2003) have justified the prominence of microbial community structures. These studies have also focused on the biogenic materials called cryoconite, which is known to consist of the organic and inorganic substances.

Some of the microbes that have been isolated from the world glaciers include the Archaea; *Euryarchaeota*, Bacteria; *Cyanobacterium*, *Cryobacterium, Agreria, Bacteriodetes*, *Actinobacteria, Acidobacteria, Chlamydiae, Chloroflexi, Firmicutes, Nitrospirae, Fibrobacteres, Plantomycetes, Proteobacteria, Spirochaeta* and fungi such as *Penicillium,* yeast, *Fussarium* and phages. Some of these microbes have got historical significance in medicine, drug production, redox potential reactions, and organic compound mineralization.

In the tropical Africa, there is high energy flow and intermittent weather conditions. The variation in weather conditions and high energy flow would be a factor to facilitate biological activities within an ecosystem, especially the undisturbed regions like the glaciers.

The surviving African glacier are located closer to the equator. Particularly, Lewis glacier is strategically located to the windward and leeward sides. Within these locations the glacier experiences seasonal snow fall, influx of the bioaerosols, upwind predominant weather systems from Indian ocean and downwind influxes from the savannah zones of Nanyuki. These conditions make Lewis glacier to have variable growth conditions to support diverse microbial communities within the cryoconite holes, which experience turbulence and inter-cryoconite interactions due to melt water runoffs and tillage to subglacial layer.

The increased multidrug resistance and biotechnological applications requires new approaches to synthesize new antibiotics from the serotypes of microorganisms. The development in industrialisation and bioengineering also requires microorganisms with

special gene products that can be utilised within the industries. Prior to the identification of the functional roles of psychrophiles, it is important to understand the microbial ecology of the rapidly melting glacier and conservation of the culturable species. This was achieved by the culture-dependent and culture-independent next generation sequencing techniques. The result from the current study is important for the creation of a local database for references to tropical glacier ecology. The findings are also important to conservation of the culturable tropical glacier isolate that can be used for subsequent analyses in the biotechnological applications.

1.4.Research Questions

1. Are there novel microorganisms in the Lewis Glacier on Mount Kenya, Nyeri County, Kenya
2. How diverse are the microorganisms in Lewis Glacier on Mount Kenya, Nyeri County, Kenya?
3. Which organisms are the indicator species on the Lewis glacier and on its foreland?
4. Is there rapid disappearing and erosion of the glacier due to climate change endangering any of the microorganisms species?
5. Are there chronological differences in microbial community structures in the Lewis Glacier and its foreland?
6. Does the environmental factors affect the microbial community structure on Lewis glacier?

1.5.Hypothesis

There is low microbial diversity and composition in the tropical Lewis glacier on Mount Kenya, Nyeri County.

1.6.General objective

To determine the metagenomes and composition of prokaryotes on Lewis glacier and its Foreland in Mount Kenya.

1.7.Specific objectives

1. To determine the composition and the abundance of prokaryotic taxa on Lewis glacier and its foreland on Mt. Kenya.
2. To determine the composition and abundance of the predominant phyla on Lewis Glacier and its foreland in Mount Kenya.
3. To determine the diversity of prokaryotes in the snow of Lewis Glacier by targeting the 16S rDNA gene.
4. To compare the distribution, composition and diversity of taxa in the Glacier and its foreland.
5. To determine the indicator species and their functional roles on Lewis glacier and its foreland.
6. To determine the relationship between the environmental factors and the microbial community structures on the Glacier surface

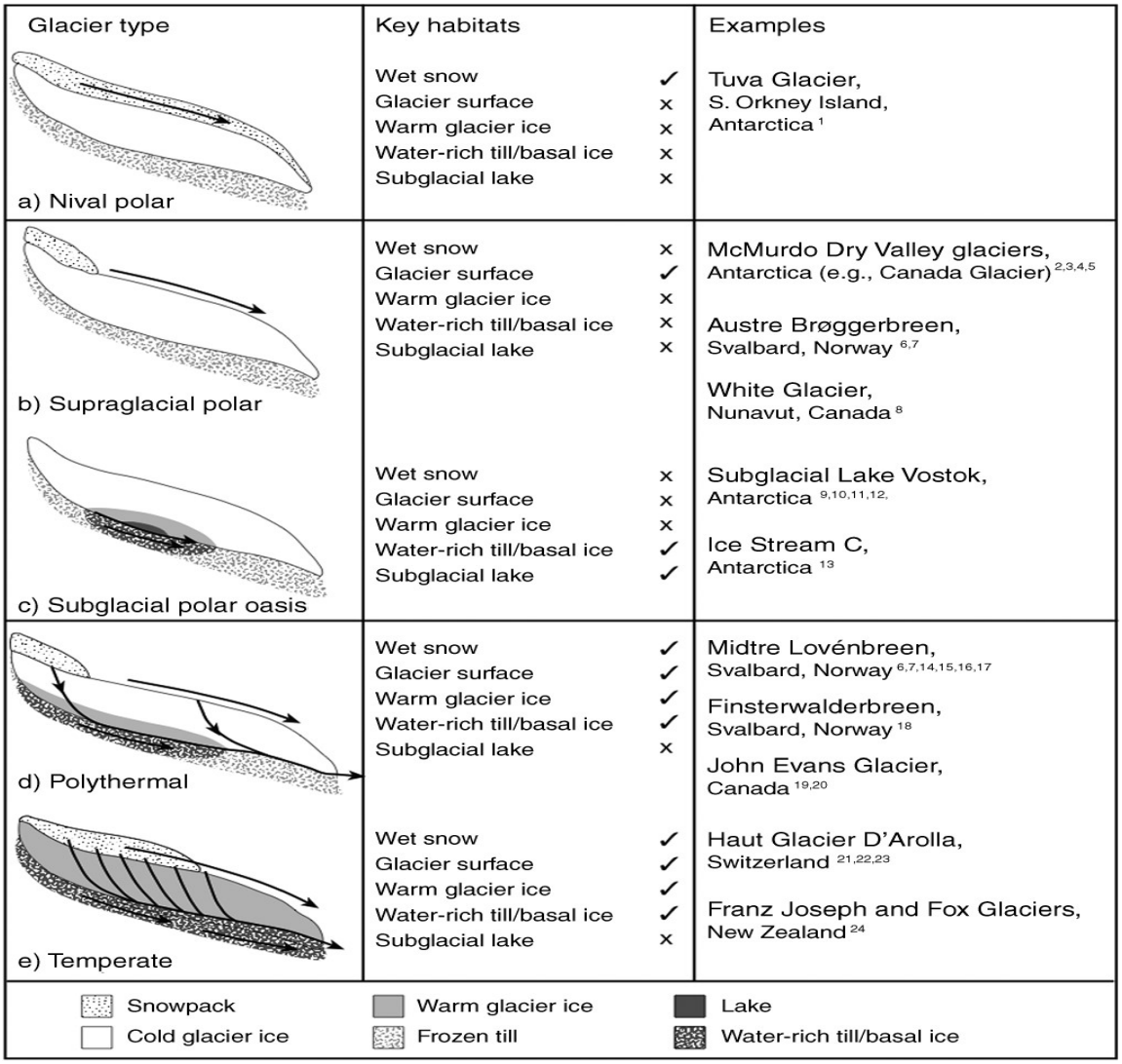
CHAPTER TWO

LITERATURE REVIEW

2.1. Biology and schematic occurrence of glacier ecosystems

Glaciers are classified as cold and temperate glaciers **(Figure 2.1a, b, c, d and e** respectively; Blatter & Hutter, 1991). Temperate glacier is characterised with ice entirely at the melting points while cold glaciers consist of the subglacial temperature variations due to microbial activities and other glacial physiological processes (Paterson, 1994; Cuffey & Paterson, 2010).

The schematic occurrence of the glacier (Cuffey & Paterson, 2010) reflects the combined surface, internal and basal heat budget on the stability of ice distribution at the pressure melting points of the ice sheets (Parizek *et al.,* 2002). Microbial community structures within the glacier ecosystems vary due to spatial biogeochemical processes within the cryoconite material and melt water (Stibal *et al.,* 2012; Telling *et al.,* 2012). The glacial structures also create habitable microenvironments for the microbial communities and their activities on the surface snow, ice and glacier surfaces (McCarroll, 2002). However, the glacial ecosystem structure variations **(Figure 2.1)** are key factors to microbial retention and loss that occur through the hydraulic melt water flow channels **(Figure 2.1).**



1, Hodson (2006); 2, Christner et al. (2003b); 3, Porazinska et al. (2002); 4, Tranter et al. (2004); 5, Wharton et al. (1981); 6, Kasˇtovska´ et al. (2005); 7, Hodson et al. (2005a); 8, Mueller and Pollard (2004); 9, Christner et al. (2001); 10, Karl et al. (1999); 11, Priscu et al. (1999); 12, Siegert et al. (2001); 13, Lanoil et al. (2004); 14, Anesio et al. (in press); 15, Mindl et al. (2007); 16, Sa¨ wstro¨ m et al. (2002); 17, Wynn et al. (2006); 18, Wadham et al. (2004); 19, Bhatia et al. (2006); 20, Skidmore et al. (2000); 21, Bottrell and Tranter (2002); 22, Sharp et al. (1999); 23, Tranter et al. (2002); 24, Foght et al. (2004).

Figure 2.1: Schematic representation of glacier structures and their most important microbial habitats. Hydrological flow of melt water is shown by the arrows

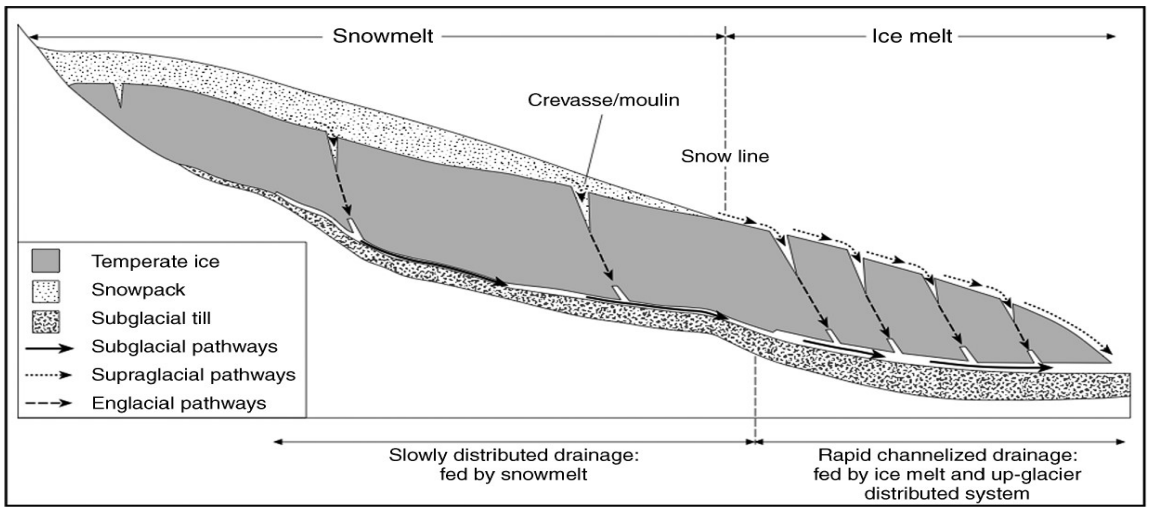


Figure 2.2: Glacier drainage system structure.

2.2. Cryoconite material in the subglacial ecosystems

Cryoconite holes are water-filled depressions on the glacial surface (Wharton *et al.,* 1985). They represent the most active microbial habitat upon snow-free, melting ice constituting the organic and inorganic particles (Takeuchi *et al.,* 2001). Cryoconite holes are common to all glacial environments where surface melting occurs (Porazinska *et al.,* 2004), but are usually restricted to the ice surface of the glacier ablation zone (Van de Wal *et al.,* 2008). The cryoconite materials melt into the ice due to its lower reflectance than the surface ice (Takeuchi *et al.,* 2001) forming a hole that becomes filled with its own meltwater **(Figure 2.2)** that may be washed into streams and melt water channels **(Figure 2.2).**

There are three types of cryoconite holes are found to exist on the glaciers: open holes (a), submerged holes (b), and closed holes (c) **(Figure 2.3).** These features are conducive microenvironments harbouring diverse microbial life forms coupled with the photoautotrophic activities and physiological processes like carbon and nitrogen cycling (Anesio *et al.,* 2017; Visser *et al.,* 2012). The microbial community structure within the cryoconite is dominated by the bacteria (mainly Cyanobacteria), algae and fungi (Kaštovská *et al.,* 2007; Säwström *et al.,* 2002). Tardigrades, rotifers, and nematodes and viruses have also been reported to exist in the cryoconite materials (Choudhari *et al.,* 2015; De Smet & Van Rompu, 1994; Säwström et al., 2002).

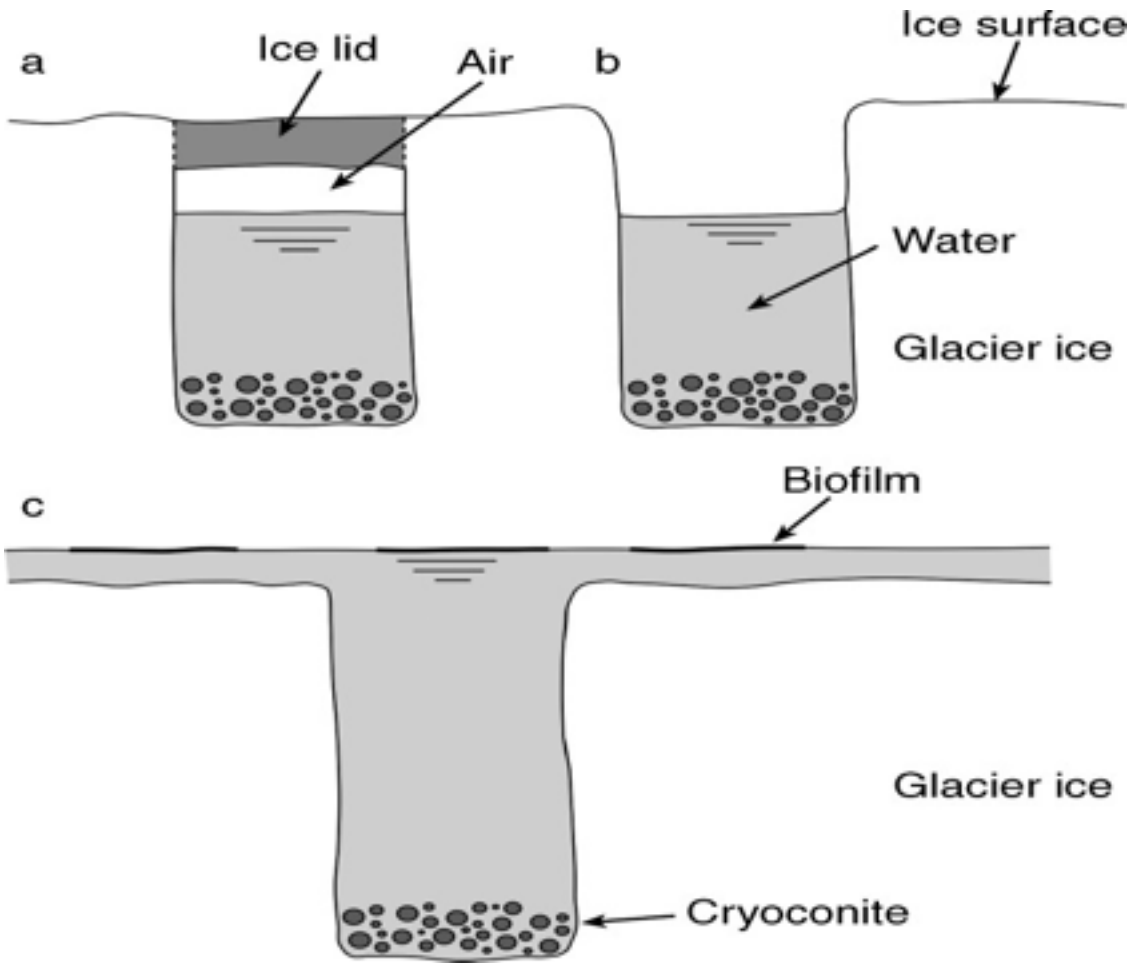


Figure 2.3: Anatomy of a cryoconite hole, showing (a) closed, (b) open, and (c) submerged morphologies.

2.3. Microbial ecology of the glaciers and ice sheets

2.3.1. Microbial ecology of the supraglacial ice cover

Polar glaciers, mid-latitude glaciers and high-latitude mountains harbour diverse community structures (Skidmore *et al.,* 2005; Palmisano & Sullivan, 1983; Grebmeier & Barry, 1991). The diversity of bacteria, fungi and algae that are present in ice sheets from the polar regions and mountain glaciers can be dated using ice cores (Christner *et al.,* 2003; (Miteva *et al.,* 2015; Uetake *et al.,* 2016).

Snow algae and yeast cells can reside and multiply on glacier surfaces from temperate regions, in accumulation areas because of the availability of meltwater, which is essential for their growth and nutrient cycling (Uetake *et al.,* 2011). Cold-adapted yeasts have been isolated from supraglacial and subglacial ice in Svalbard, Norway (Butinar *et al.,* 2007), Austrian glacier ice (Margesin *et al.,* 2003), Italian subglacial meltwater (Branda *et al.,* 2010b), supraglacial and subglacial ice and meltwater from the Italian Alps (Turchetti *et al.,* 2008), glacial and subglacial waters from northwest Patagonia (Brizzio *et al.,* 2007; De García *et al.,* 2007), and an Antarctic deep ice core (Amato *et al.,* 2009).

Temperate glaciers are characterised with the seasonal snow cover that harbours phytoflagellates, mostly the micro-algae (Anesio *et al.,* 2017). Snow-packs are good sources of microbial inoculi (Hallbeck, 2009), nutrients (Schmidt & Lipson, 2004), and water (Barnett *et al.,* 2005) that cascades through the drainage tills forming the melt water at the subglacial ecosystems (**Figure 2.1**). The availability of the inorganic nitrogen and phosphate compounds (Uetake *et al.,* 2011) on glacier environments is controlled by the phototrophic activity of the snow algae and the snow-pack *Cyanobacteria* (Hodson *et al.,* 2008a). The 16S rRNA gene sequencing has also revealed the availability of *Proteobacteria, Firmicutes* and *Actinobacteria* that are able to degrade organic compounds such as the propionate, acetate and formate that are available on the Arctic snow (Hodson *et al.,* 2008a). However, the phototrophic and heterotrophic impacts of these microbes on the glacier snow-packs are not well understood.

2.3.2. Microbial ecology of the supraglacial ice

Algae and the phytoflagellates also colonise the supraglacial ice within the cryoconite holes (Hodson *et al.,* 2008a) and even in the distinct supraglacial ice mats (Vincent *et al.,* 2004). The ice mats within the supraglacial regions also support photosynthetic processes making it as productive as those of the marine ecosystems (Vincent *et al.,* 2000).

2.3.3. Microbial ecology of the cryoconite holes

Literature indicates that the diversity of microorganisms in the cryoconite holes is much greater than that of the ice cover and the mats (Takeuchi *et al.,* 2000; Porazinska *et al.,* 2002; Säwström *et al.,* 2002; Christner *et al.,* 2003; Kaštovská *et al.,* 2005; Hodson *et al.,* 2008). However, the variability of the microbial community structure is dependent on the environmental conditions; especially the glacier slopes and inter-cryoconite hole mixing that leads to continuous colony flushing by melt water and ecosystem homogenization, respectively (Hodson *et al.,* 2008).

Cryoconite holes from the polar glaciers have been reported to harbour filamentous fungi and yeast (Margesin *et al.,* 2003). However, most cryoconite are known to be dominated by the *Cyanobacteria,* phytoflagallate bacteria, algae and other fungi (Kaštovská *et al.,* 2005; Hodson *et al.,* 2008). Viruses are no exception to the cryoconite regions (Säwström *et al.,* 2002), they have been isolated from the melt water, cryoconite granules and sediments (MacDonell & Fitzsimons, 2008). The role of viruses on glacier is highly significant as they control the bacterial and fungal availability through cell lysis of the hosts (Hodson *et al.,* 2008). The viral activity, however, requires further investigation to understand their role and impact on the glacier ecosystem.

2.4. Microbial community structure in the glacier foreland

The response of glacier thermal regime to climate change is a critical phenomenon upon ecosystems in supraglacial and subglacial environments (Hodson *et al.,* 2008). Microbial activity within the cryoconite holes lead to the production of cohesive compounds, called extracellular polymeric substances, which cause the aggregation of cryoconite on the ice surface (Takeuchi *et al.,* 2001; Hodson *et al.,* 2010; Langford *et al.,* 2010). Larger compound aggregates may have a longer residence time on the ice and thus enhance surface melting by their prolonged reduction of surface albedo (Irvine-Fynn *et al.,* 2011).

The large aggregates may also provide more conducive microenvironments for a wider range of microorganisms and physiological processes than smaller aggregates due to diffusion of oxygenated surface waters by the redox conditions found there (Hodson *et al.,* 2010; Langford *et al.,* 2010). These factors then enable cryoconite to contribute to ecological succession in the glacier fore field following the continued retreat of the glacier (Kaštovská *et al.,* 2007; Edwards *et al.,* 2013).

The foreland of a receding glacier is extremely heterogeneous in terms of physical landforms, soil structure and environmental conditions, each of which directly impact the composition, activity and function of the microbial community (Brankatschk *et al.,* 2011; Knelman *et al.,* 2012; Zumsteg *et al.,* 2012). The biogeochemical process of the soil is not just the result of a single microbe, but the whole microbial community (Bernasconi, 2008; Knelman *et al.,* 2012). DNA extraction and amplification has revealed the difference in microbial community composition in glacial foreland soils (Zumsteg *et al.,* 2012). Some of the microorganisms that have been isolated from the glacier forefield include Cyanobacteria, Proteobacteria, Actinobacteria, Archaea and various species of fungi from the Arctic and Antarctic (Sigler *et al.,* 2002; Zumsteg *et al.,* 2012). However, the microbial diversity and functional status of the tropical glacier foreland remain unknown.

2.5. Phylogenetic and functional diversity of microorganisms on the glacier

The biogeographical and microbial community structures from the cold environment have been a major concern to microbial ecologists and molecular biologists. Microbial diversity and their phylogenetic relationships have been achieved based on the culture dependent methods and culture independent methods. Initially, the glacial microbial diversity studies were based on the carbon dating methods, ^{14}C-labeled organic compounds (Bardgett *et al.,* 2007). Through these techniques, both the aerobic and anaerobic microbes were found (Skidmore *et al.,* 2000) to colonize the supraglacial ice sheets and the subglacial cryoconite materials, respectively (Stibal *et al.,* 2006; Kaštovská *et al.,* 2007).

Functionally, aerobic chemo-heterotrophs, anaerobic nitrates, sulphate reducing bacteria, methanogens and ammonia-oxidizing archaea have been identified from the glacial ecosystems (Skidmore *et al.,* 2000; Wadham *et al.,* 2004; Tung *et al.,* 2005; Zarsky *et al.,* 2013). The accumulation of surface debris to melt glaciers in the cryoconite creates a microenvironment for the fungi, bacteria, viruses and archaeal species resulting to active metabolic processes (Foreman *et al.,* 2007; Edwards *et al.,* 2013). Theses metabolic processes are however, linked to the accelerated ice melting (Kohler *et al.,* 2007; Gardner *et al.,* 2011) resulting to considerable amount of carbon accumulations (Slater *et al.,* 2002; McConnell *et al.,* 2007) and nutrient cycling (Björkman *et al.,* 2013).

Active microbes inhabiting the cryoconite materials produce cohesive extracellular polymeric compounds that cause cryoconite aggregation on the ice surface (Hodson *et al*., 2010; Langford *et al*., 2010). The aggregates may have prolonged residence period on the ice surface leading to reduction of the surface albedo that enhances ice melting (Irvine-Fynn *et al.,* 2011). The aggregates are also coupled with the diverse microbial communities that are associated with potential redox conditions (Hodson *et al.,* 2010, Langford *et al.,* 2010). These factors have significantly contributed to the ecological successions in the glacier foreland due to sudden glacier retreats (Edwards *et al.,* 2013). Multi-cluster analysis of the geochemical processes, surface debris and meltwater can be used to establish functional diversity of microorganisms on the glacier surfaces. The functional diversity of microorganisms can also be established by the PCR amplification of the 16S rDNA, genes followed by the pyrosequencing and multivariate statistical analysis to better understand microbial carbon and nitrogen cycling within the glacier ecosystem. Microbial phylogenetic diversity has also been analysed by the amplification of the 16S rDNA targeting the V3, V4 and V7 regions for the prokaryotes (Zarsky *et al.,* 2013).

2.6. Albedo reduction

The glacial surface albedo is the name given to the light reflectance of snow and ice (Takeuchi *et al.,* 2001). This is an important phenomenon because the light reflectance affects glacial melting (Thomas and Duval, 1995; Hoham and Duval, 2001). Some glaciers in Himalayas are covered with a large amount of cryoconite that comes from snow algae and bacteria derived from snow algae and bacteria (Kohshima *et al.,* 1993; Takeuchi et al., 2001). Usually, the melting rates of the intact surfaces are reported to be three times larger than that of the surfaces without the cryoconite (Takeuchi *et al.,* 2001). Thus, microbial activity on the glacier possibly affects heat budget and mass balance of glaciers leading to rapid deglaciation.

2.7. Mount Kenya Lewis Glacier

Lewis glacier lies astride the equator. In 1993, it was reported to be one of the smallest glacier (0.4 km^2) in Africa (Kaser, 1999) and the biggest glacier in Mt. Kenya (Lewis Glacier), which is now rapidly shrinking (Prinz *et al.,* 2011). The size and shape of the surviving African glaciers keep on changing. Probably this is based on the endogenous and exogenous glacier experiences. Notably, cloud cover, sunlight radiation and atmospheric humidity has been suggested to significantly contribute to the (Mölg *et al.,* 2009).

The structure of Lewis glacier in to two sites with different slope angles is of concern to the rapid glacier loss (Thomson *et al.,* 2009; Rabatel *et al.,* 2013). The abundant cryoconite materials and surface dust particles (Kuja *et al.,* 2018) are a characteristic of rich inocula harbouring complex extracellular polymeric substances with diverse microbial community structure as described elsewhere (Hodson *et al.,* 2008). The complex association within the cryoconite holes attributed by the organic and inorganic particles can attribute to biogeochemical processes (Hodson *et al.,* 2008; Stibal *et al*., 2012).

This study employed space-for-time substitutions with age, such that distance from the glacier terminal was used as a proxy for soil age soils being further from the glacier's edge. The majority of foreland studies have focused on low and high cover plant succession (Bormann & Sidle 1990; Matthews, 1992; Chapin *et al*., 1994; Mizuno and Fujita, 2014) and, while vegetation succession can be stochastic across different environments, research has shown that some consistent patterns of plant-mediated changes to ecosystem functioning such as nitrogen (N) fixation (Walker & del Moral, 2003) are absolutely significant.

For instance, plant communities dominated by N-fixing species are common early in the primary succession of newly deglaciated soils in Glacier Bay, Alaska (Reiners *et al.,* 1971). These N-fixing plants increase soil N pools and N-cycling rates (Bormann & Sidle 1990; Chapin *et al.,*1994) and affect soil development (Crocker & Major, 1955). Such plant-driven changes to soil N cycling have significant effects on ecosystem functioning and on the establishment of subsequent plant communities (Chapin *et al.,* 1994). The application of modern molecular approaches, like the next generation sequencing, it is becoming clear that microbes rapidly colonize deglaciated soils long before vascular plants or lichens appear (Jumpponen 2003; Tscherko *et al.,* 2003; Nicol *et al.,* 2005; Bardgett *et al.,* 2007).

CHAPTER THREE

MATERIALS AND METHODS

3.1. Study site

Study site was Mt. Kenya Lewis glacier (latitude 0° 9′ 30″S - 0° 9′ 15″S, longitude 37° 18′ 45″E - 37° 19′ 0″E, Fig. 3.1). Lewis glacier on Mt. Kenya is the smallest (0.4 km^2 in 1993) in Africa (Kaser, 1999) and biggest glacier on Mt. Kenya (Lewis Glacier) that is rapidly shrinking (Prinz *et al.,* 2011). All necessary collection permits were obtained via National Commission for Science, Technology and Innovation (NACOSTI), Kenya Wildlife Service (KWS) and National Environmental Management Authority (NEMA).

3.2. Sampling from the supraglacial surface and foreland

Five sites were identified based on the stakes located on the glacier prior to sampling on the glacier and GPS locations indicating the last glacier terminal in the foreland sites. The sampling stakes were 10 cm cubed while the distance between the stakes were 30 cm. The glacier site has the upper less sloppy (17.4 degrees, site 1) and the sloppier section (19.8 degrees, site 2). The foreland was structured into side foreland (SF) and glacier foreland (GF). Cryoconite materials were collected from the glacier while soil samples were collected from the foreland. Both samples were collected by scooping using sterile steel shovel and poured into the sterile plastic bags for laboratory isolation. The collected were stored in the cool box containing ice in the field and later on shipped to Tokyo, National Institute of Polar Research laboratories under dry ice preservation for laboratory analysis.

3.3 Microscopic observations of microorganisms

A drop of melted snow and ice samples were poured onto slides and covered for observation light microscope at X40 magnification. Shapes and pattern of the cells were observed and recorded for characterization.

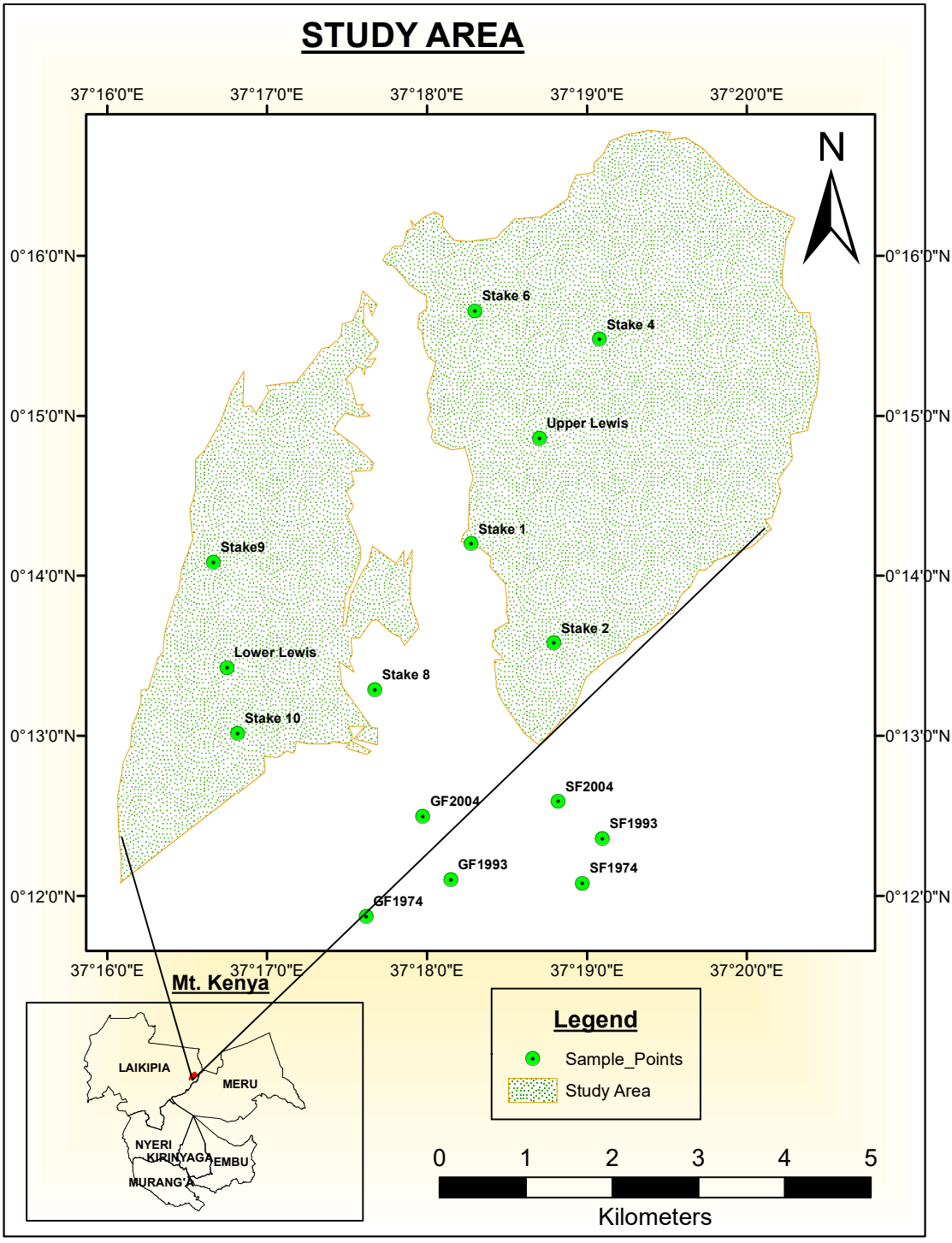


Figure 3.1: Map of Lewis glacier showing the glacier study area and sampling sites.

3.4. Isolation and Physiological characterization of the isolates

Sterile syringe were used to inoculate 200 *µl* of melt ice and dissolved soil suspensions from the foreland into plates containing appropriate bacterial media (R2A). For *Cyanobacteria* isolates, BG-11 broth was used to inoculate melt water into a hundred folds. Inoculants were incubated at 4°C, 10°C, and 25°C. Growth rates were monitored daily while sub-culturing colonies into independent plates until axenic cultures were obtained. All cultures were treated as glacier isolates and were independently analysed. All the isolation procedures were carried out within a class 100 laminar flow clean bench to avoid contamination.

3.5. Molecular characterization

3.5.1 DNA extraction and purification

Total microbial community DNA were extracted from all the samples; pellets from melt water samples, 0.25g of sediment samples and 0.4g of foreland soil samples using the phenol chloroform method (Sambrook *et al*., 1989). This was used as template for the subsequent Illumina sequencing.

For the culture-dependent methods, pure isolates were picked from individual sterile plates using a sterile wire-loop and directly added to a mixture of MP FastDNA soil kit (FastDNA™ Spin Kit for Soil) protocol. Purified DNA was quantified photometrically (NanoDrop; Thermo Fisher Scientific, Germany) and used as a template for amplification of 16S rDNA genes using the bacterial primer pair 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'- GGTTACCTTGTTACGACTT-3') (Lane, 1991). TaKaRa Ex Taq DNA polymerase (Takara, Shiga, Japan) was used for the PCR amplification. For each PCR, 1 µl of the template was mixed with TaKaRa Ex Taq™ HS (5 units/µl) according to the protocol. The PCR conditions were as described by Mackenzie et al. (2007). PCR product size was checked using a 1.5 % agarose gel stained with ethidium bromide. The amplicons were gel purified using Macherey-Nagel

NucleoSpin extract II kit (NucleoSpin® kits on QIAcube®) and eluted in 30 µl of TE Buffer (5 mM, pH 8.0).

3.5.2 Amplicon library preparation and sequencing for culture independent

PCR amplification of the V3/V4 variable region of the 16S rDNA gene was carried out from the extracted samples DNA, using bacteria primers 16S F for Miseq (TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG) and 16S R for Miseq (GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTATCC) according to Caporaso *et al.* (2012). The primers were further treated with overhang adaptors at the locus-specific primer for the V3/V4 region as Forward overhang (TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG) and Reverse overhang: (GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG). The reaction volume (25 µl) contained 2.5 µl of microbial DNA (5 ng/µl), 5 µl of forward primer (1 µM), 5 µl of reverse primer (1 µM) and 12.5 µl of HotStarTaq Plus Master Mix Kit.

Polymerase Chain Reactions proceeded in a 35 cycles using the HotStarTaq Plus Master Mix Kit (Qiagen, USA) with initial denaturation heating at 95°C for 3 minutes, followed by 30 cycles of denaturation at 94°C for 30 seconds, annealing at 54°C for 40 seconds and extension at 72°C for 1 minutes, and a final elongation at 72°C for 5 min. This was then followed by an amplicon PCR where the template was amplified at the region based on the interest specific primers. The denaturation for the amplicon template was at 95ºC for 3 minutes, followed by 25 cycles 95ºC for 30 seconds, 55ºC for 30 seconds, 72ºC for 30 seconds, an extension at 72ºC for 5 minutes and the amplicon was finally held at 4ºC.

The quality of PCR products was assessed on 2 % agarose gel to determine the success of amplification and the relative intensity of bands. Multiple samples, tagged with different barcodes, were pooled in equimolar ratios based on their DNA concentrations from the gel images. Pooled samples were purified using calibrated Ampure XP beads (Beckman Coulter) for use in library preparation. The pooled and purified PCR products were used

to prepare 16S rDNA library by following Illumina TruSeq DNA library preparation protocol (Yu and Zhan, 2012). PCR products were sequenced with a 3130xl Genetic Analyzer (Applied Biosystems, California) at the National Institute of Polar Research (NIPR).

3.5.3 Sequence analysis and phylogenetic classification

Raw sequencing reads were quality trimmed according to published recommendations (Huse *et al.,* 2007) using the QIME2 pipeline version 2018.11.0. (Allali *et al.,* 2017). The denoised sequences were then evaluated for potential chimeric sequences using UCHIME within the USEARCH package v.4.2.66 (Edgar, 2010). A sequence identity cutoff of 97% were used to pick Amplicon Sequence Variants (ASVs) from the quality filtered non-chimeric sequences. Representative ASVs were picked using QIIME2 (Allali *et al.,* 2017). Taxonomies were assigned from the GG reference library using a BLAST method (Altschul *et al,.* 1990) with removal of shared gaps before an ASV table is made at dissimilarity levels of 3%, 5% and 10%. Rarefaction curves and diversity indices were calculated and plotted for each sample using QIIME2 (Allali *et al.,* 2017). Principal coordinate analysis (PCoA) of weighted and unweighted UniFrac were also calculated and plotted with QIIME (Lozupone and Knight, 2005; Allali *et al.,* 2017). ACE and Chao1 (Chao and Bunge, 2002) indices were calculated using the EstimateS program version 8.2.0 (http://purl.oclc.org/estimates) and normality test repeated. Principal Components Analysis (PCA) was performed by using R program version 3.5.1 (ter Braak and Schaffers, 2002) in order to compare bacterial community structures across all samples based on the relative abundance of bacterial phyla and *Proteobacteria* classes. Amplicon data sets were submitted to the National Center for Biotechnology Information (NCBI) Sequence Read Archives.

3.5. Analysis of total organic matter, albedo and nutrient concentration

Albedo measurements were taken from the field using albedometer. The concentration of dissolved organic carbon (DOC) was measured in YEPD agar medium diluted 500-fold

(1/ 500 YEPD) and in the nutrient eluted supernatant of the 10% UWA medium using a Spectroquant test kit (Merck, Darmstadt, Germany) and a spectral photometer (photo-Lab S12; Wissenschaftlich-Technische Werkstätten, Weilheim, Germany). All samples and media were filtered through a polycarbonate membrane (K040A047A; Advantec, CA). The nutrient eluted supernatant of the 10% UWA agar medium were collected after shaking (110 r.p. m.) for 24 h, according to Uzuka (1992). The melted surface-ice samples were mixed from two aliquots as described by Uetake *et al.* (2012).

3.6. Statistical analysis

R version 3.5.2 was used for all the analyses. Diversity indices (Richness, Shannon, InvSimpson, absolute diversity), rarefaction curves and Venn diagram (to compare the shared ASVs between the glacier and foreland sampling sites) were calculated from the resulting ASVs using Vegan package version 1.16–32.. Hundred iterations of rarefaction were computed for each sample to 20,000 sequence using QIME2 2018.11.0. (Allali *et al.,* 2017). Chao1, Shannon and InvSimpson diversity indices, non-parametric estimation of ASV richness between the data sets calculated from the rarefied data was used to compare species richness between the data sets, sample sites and sample types. Indicator species analysis was performed using the *labdsv* function implemented in the *IndVal* package.

Community and Environmental distances were compared using Analysis of similarity (ANOSIM) and adosin function (PERMANOVA) test, based upon Bray-Curtis dissimilarity and Euclidean distance measurements with 999 permutations. Significance was tested at 95% confidence interval ($p = 0.05$). Non-metric Multidimensional Scaling (NMDS), distance based Redundancy analysis (dbRDA), component analyses of proximity (CAP), as well as Hierarchical clustering of the environmental data, and the community data, based on Bray-Curtis dissimilarity, were carried out using the R programming language (Team, 2014) and the Vegan package (Oksanen *et al.,* 2011). Distance based Redundancy analysis (RDA) was used in order to assess the relationships

between known environmental variables and variation in the multivariate data based on the Euclidian. 999 permutations in the unrestricted mode, and manual forward selection, were used for the RDA. The results of the dbRDA were summarized using biplot diagrams.

Correlation, based on Spearman, Pearson's and Kendall correlation coefficient, analysis between the environmental samples and the community structure was conducted and significance was tested using Mantel test (Bonnet & Van de Peer, 2012). To support ASV-based analysis, taxonomic groups were derived from the reads assigned to each taxon at all ranks using the taxa_summary.txt output from QIIME2 pipeline Version 2018.11.0. Distances from the glacier terminus were treated either as a continuous (linear regression) or categorical variable (ANOVA) and ages were calculated based on the last glacier terminal observed by the time of sampling in 2016, the earliest was treated at 12-years-old (2004), mid-age as 23-years-old (1993) and the oldest age as 42-years-old (1974).

CHAPTER FOUR

RESULTS

4.1 Composition and abundance of prokaryotic taxa in Lewis glacier and its Foreland

The analyses of 16S rDNA revealed the abundance of 12 phyla across the glacier and its foreland. *Cyanobacteria* was the most prominent phylum in the glacier samples with a relative abundance of 37% in both the upper and lower glacier sites. The abundance of *Cyanobacteria* decreased along the glacier foreland with the increase of age from the primary foreland to old soil as in **Figure 4.1.** *Cyanobacteria* was also evident in the newly deglaciated foreland soil (SF1; **Figure 4.1**). Its abundance in the rhizosphere samples was, however, less than 1% and could not be quantified. *Proteobacteria* was the second most prominent phyla in the glacier samples and the most prominent in the foreland samples.

The relative abundance of *Proteobacteria* ranged between 19 % - 22 % in the glacier samples, 20% - 41% in the side foreland, and 29% - 34% in the glacier foreland (**Figure 4.1**). *Bacteroidetes* was the third most abundant phyla on the glacier with the relative abundance of 15% on the upper glacier and 14% on the lower glacier, whereas its abundant ranged between (4% - 14%) across the side foreland samples and between (9% - 17%) in the glacier foreland samples (**Figure 4.1**). Phylum *Actinobacteria* was the fourth most abundant with a relative abundance of 12% on the upper and 9%) on the lower glacier, it's abundance ranged between (14% - 38%) in the side foreland samples and between (12% - 30%) across the glacier foreland samples (**Figure 4.1**). Other 8 phyla were equally distributed across the sample types with relatively overlapping abundance.

The clustering between samples collected from the glacier were closer together and the samples from the side foreland and glacier foreland were also clustered closer together (**Figure 4.2**). Soil samples were clustered together as well as the rhizosphere samples (**Figure 4.2**). The composition of taxa in the newly deglaciated foreland is further supported by the hierarchical clustering in which, SF1 was clustered together with the glacier samples. Phylum *Cyanobacteria* was clustered separately from other prominent

phyla as it occurred to dominate glacier samples The other three major phyla, *Proteobacteria, Bacteriodetes* and *Actinobacteria* were hierarchically clustered together just after the phylum *Cyanobacteria* (**Figure 4.2**).

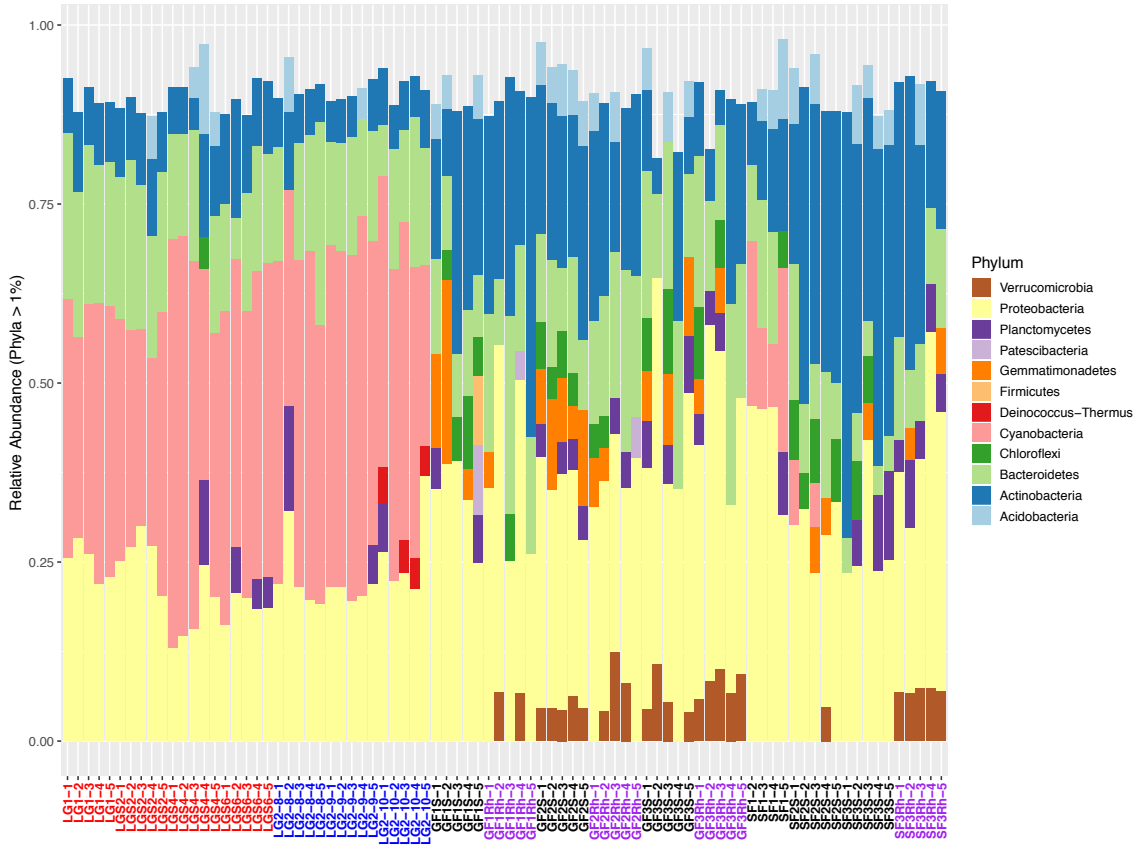


Figure 4.1: Prokaryotic taxa at genus level in samples from the Lewis glacier and its foreland. LG1 = Upper Lewis site (red), LG2 = Lower Lewis site (blue), GF = glacier foreland, Rh = rhizosphere (purple), Soil = black, SF = Side foreland

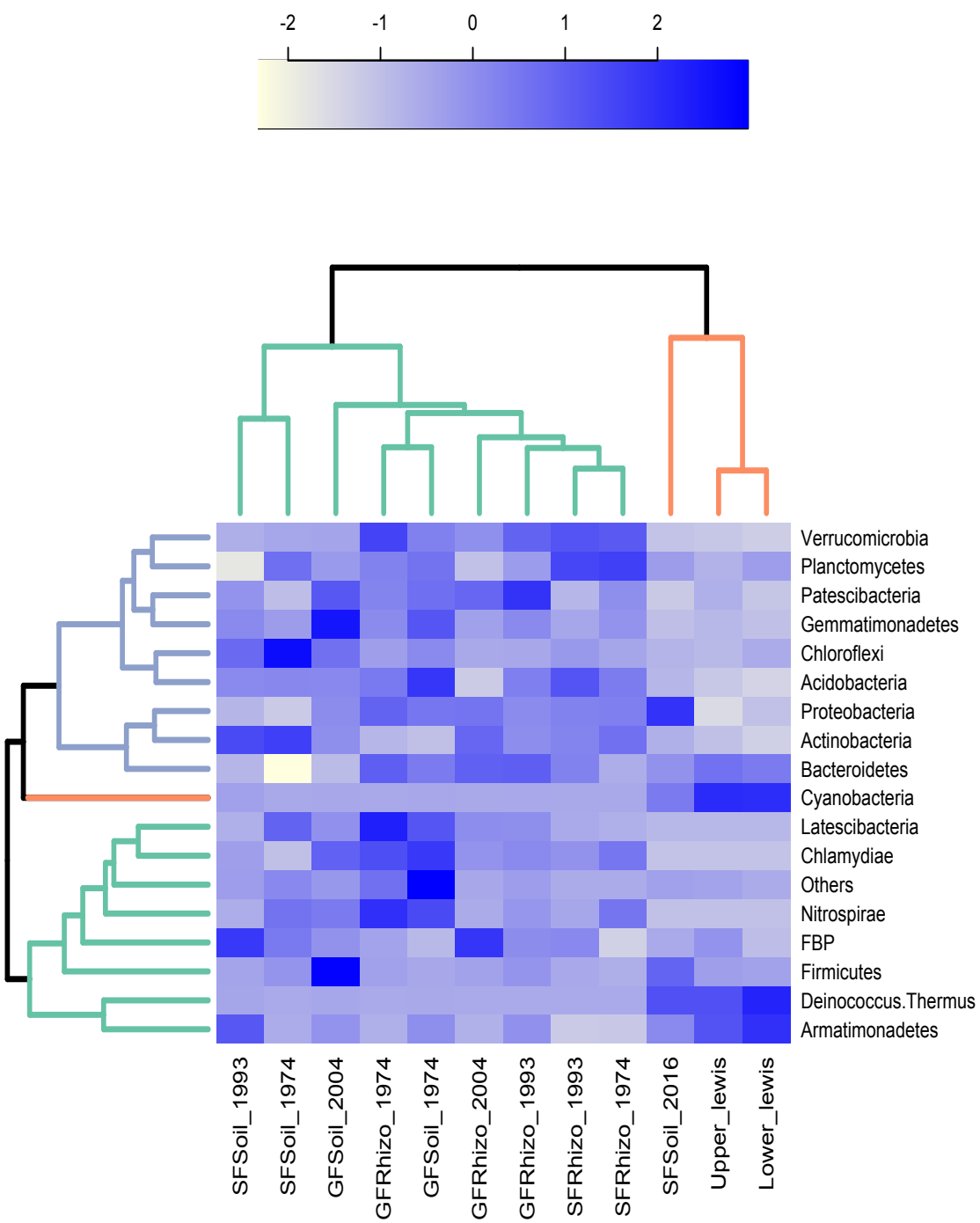


Figure 4.2: Hierarchical clustering of phyla to assess the relationships between samples and taxa. UL = Upper Lewis, LL = Lower Lewis, SF = Side foreland, GF = glacier foreland.

4.2 Composition and abundance of the predominant phyla

4.2.1 Abundance of *Cyanobacteria* in the Glacier.

The ASVs were distributed among 4 major genera of the Phylum *Cyanobacteria.* Genus *Tychonema* was distributed with a percentage between 2 % – 85 %. It was highly abundant across the upper glacier samples than the lower glacier samples. Genus *Phormidesmis* occurred with a relative percentage of 1 % – 38 % across the upper and lower glacier samples. *Phormidesmis* was more prominent in the upper glacier sites (S1LG1 and S2LG1) than the lower sites (**Figure 4.3)**. Genus *Chamaesiphon* occurred with a percentage of 1 % – 35 % across the lower and upper glacier samples respectively. It more prominent on the upper glacier site, S1LG1, S2LG1, S4LG1 and S6LG1 (**Figure 4.3**). Genus *Calothrix* was only present in upper site of the glacier with the lowest relative abundance (**Figure 4.3**)

Hierarchical clustering, based on Bray-Curtis dissimilarity, of the abundant *Cyanobacteria* at genus level in the glacier samples revealed that, the composition of the *Cyanobacteria* genera clustering is based on locations of the stakes on either the upper or lower glacier sites (**Figure 4.4**). Genera *Phormidesmis, Chamaesiphon, Pseudanabaena* and *Leptolyngbya* had high community structure at the upper glacier samples, genus *Tychonema* had a relative even distribution across the upper and lower glacier samples, genus *Oxyphotobacteria* was out-grouped in the cluster (**Figure 4.4**). Samples LG1.4.4 and LG2.8.2 are upper and lower glacier samples respectively that were outstanding in the heat map clusters due to low genera diversity (**Figure 4.4**)

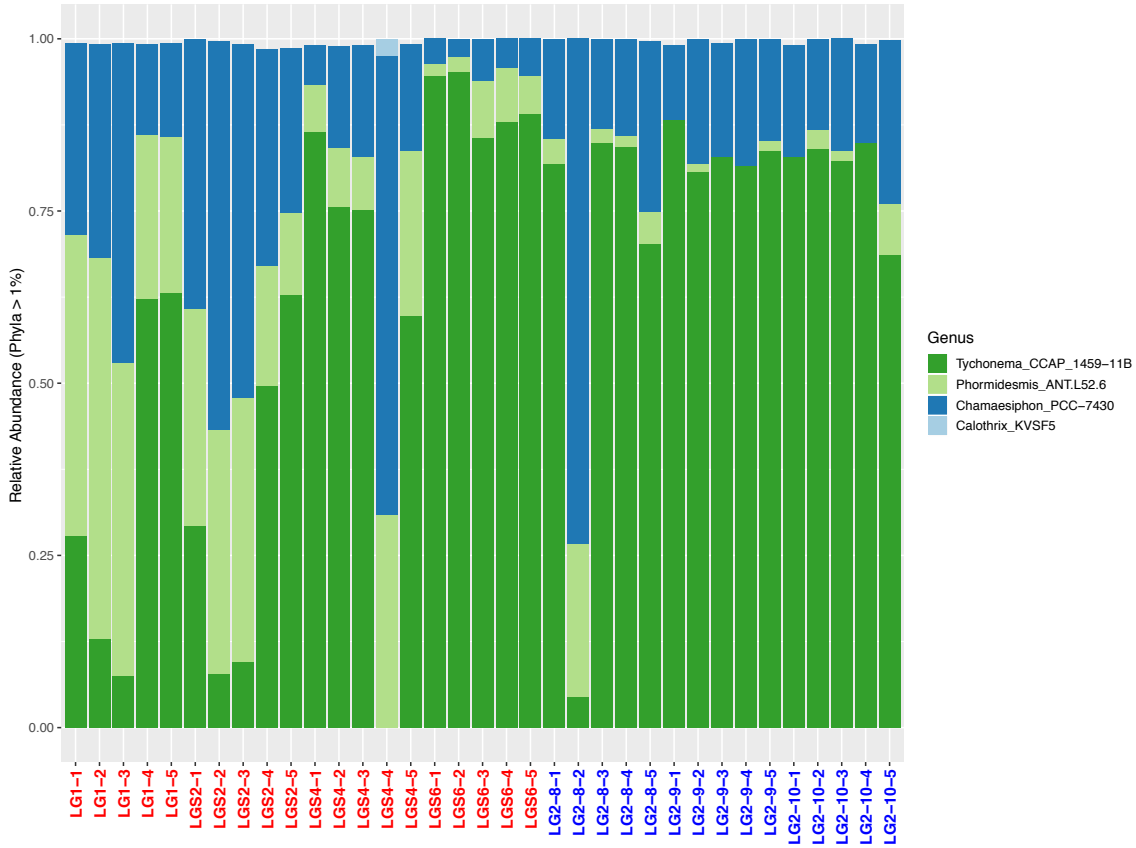


Figure 4.3: Genera of *Cyanobacteria* on the glacier samples. LG1 = Upper Lewis site (red), LG2 = Lower Lewis site (blue).

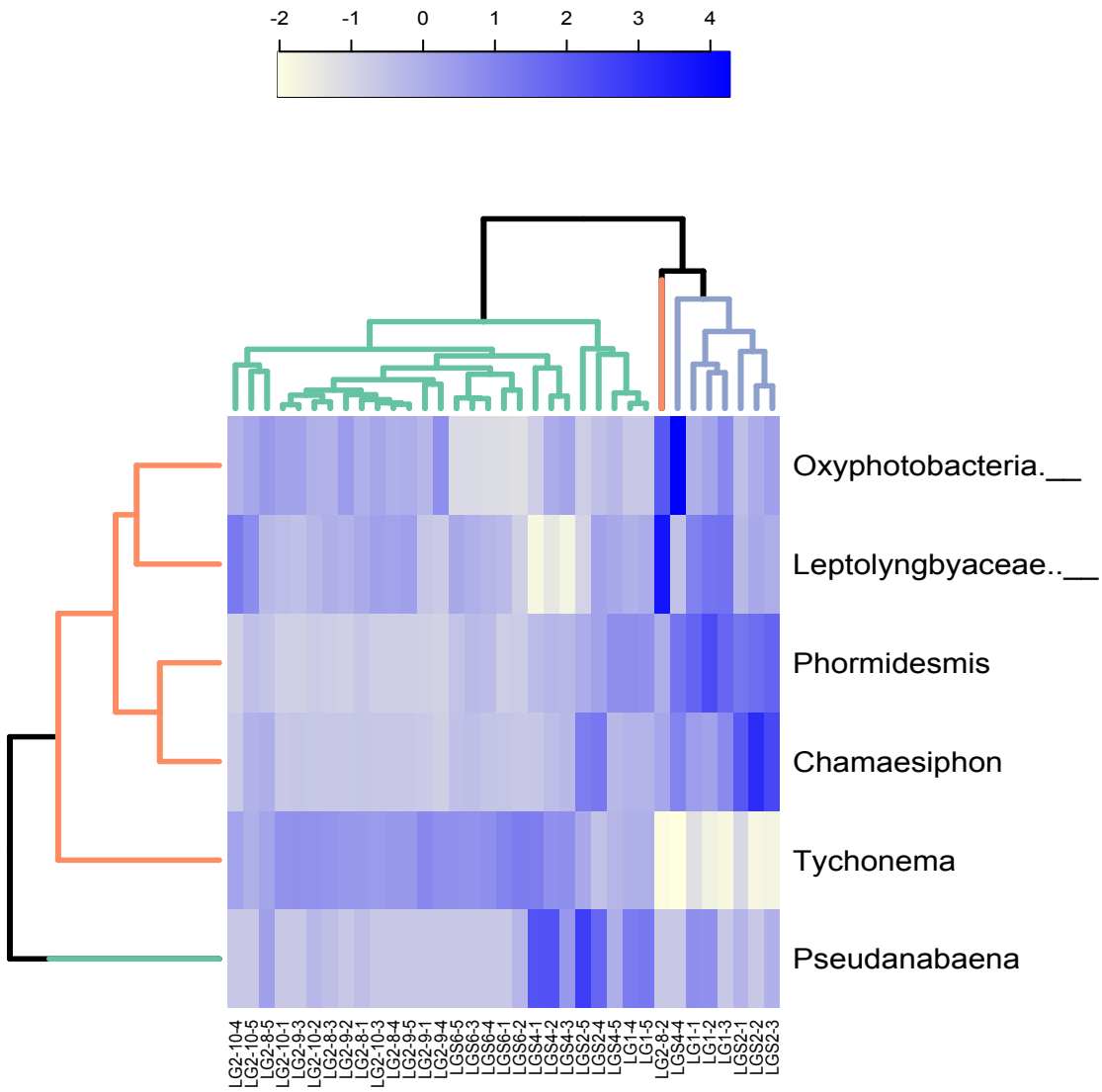


Figure 4.4: Hierarchical clustering of genera of *Cyanobacteria* in glacier samples. LG = Lewis glacier.

4.2.2 Abundance of the phylum *Cyanobacteria* in Lewis Glacier foreland.

At the foreland, the ASVs were distributed among 9 genera of the Phylum *Cyanobacteria* across the primary foreland soil, old soil and the plant rhizosphere samples **(Figure 4.5)**. The primary deglaciated soil and rhizosphere samples had diverse community structures than the old soil and rhizosphere samples along the foreland chrono-sequence. Genus *Tychonema* dominated the side-foreland soil samples and a few rhizosphere samples with a relative abundance of between 0.4 % - 54 %. In the side-foreland, only genera *Tychonema* and *Crinalium* dominated the old soil sample with a relative abundance of 54 % and 100 % respectively. Genus *Chaemasiphon* occurred in the old side-foreland rhizosphere (GF3S and GF1S) with the relative abundance of more than 90 % while in the glacier foreland it was prominent in the 12-year-old rhizosphere (GF2Rh) sample. Genus *Phormidesmis* colonized the 12-year-old rhizosphere (GF1Rh), 12-year-old foreland soil (GF2S and SF2S; 15 % - 66 %; **Figure 4.5**).

The relative abundance was also supported by the hierarchical clustering, based on Bray-Curtis dissimilarity, of the abundant *Cyanobacteria* at genus level in the foreland samples. The cluster revealed that, the composition of the *Cyanobacteria* genera is based on locations and age of the sample types at the deglaciated foreland (**Figure 4.6**). The heat map depicts the hierarchical clustering of the major *Cyanobacteria* genera. It shows a distinct pattern from the 12-year-old foreland soil, 23-year-old soil and rhizosphere, to the 42-year-old foreland soil and rhizosphere samples (**Figure 4.6**).

The prominent genera *Leptolyngbya*, *Chamaesiphon,* and *Phormidesmis* were clustered together at the primary foreland soil (12-year-old foreland soil). Genus *Chamaesiphon* was however, hierarchically clustered along the foreland chrono-sequence from the 12-year-old foreland soils, 23-year-old soil and rhizosphere, to the 42-year-old foreland soil and rhizosphere samples (**Figure 4.6**). Genus *Crinalum* distinctively clustered to the 23-year-old soil sample while genus *Tychonema* clustered to 12 and 23-year-old foreland soil samples (**Figure 4.6**). Heat map, **Figure 4.6**, showed that primary foreland soil had high ASV diversity, that steadily decreased down the foreland chrono-sequence. The samples

were also clustered according to the sample type and age along the foreland chrono-sequence.

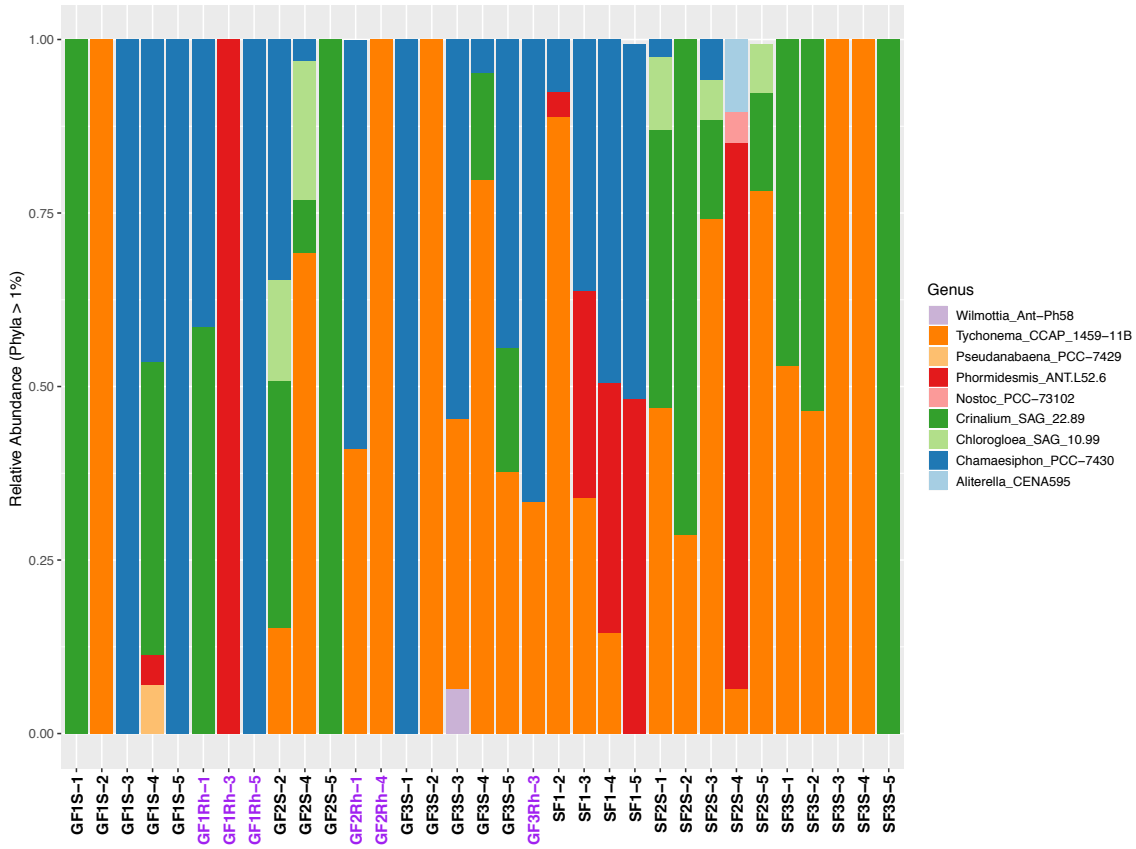


Figure 4.5: Genera of *Cyanobacteria* in the foreland sample means. GF = glacier foreland, Rh = rhizosphere (purple), SF = Side foreland, soil = black.

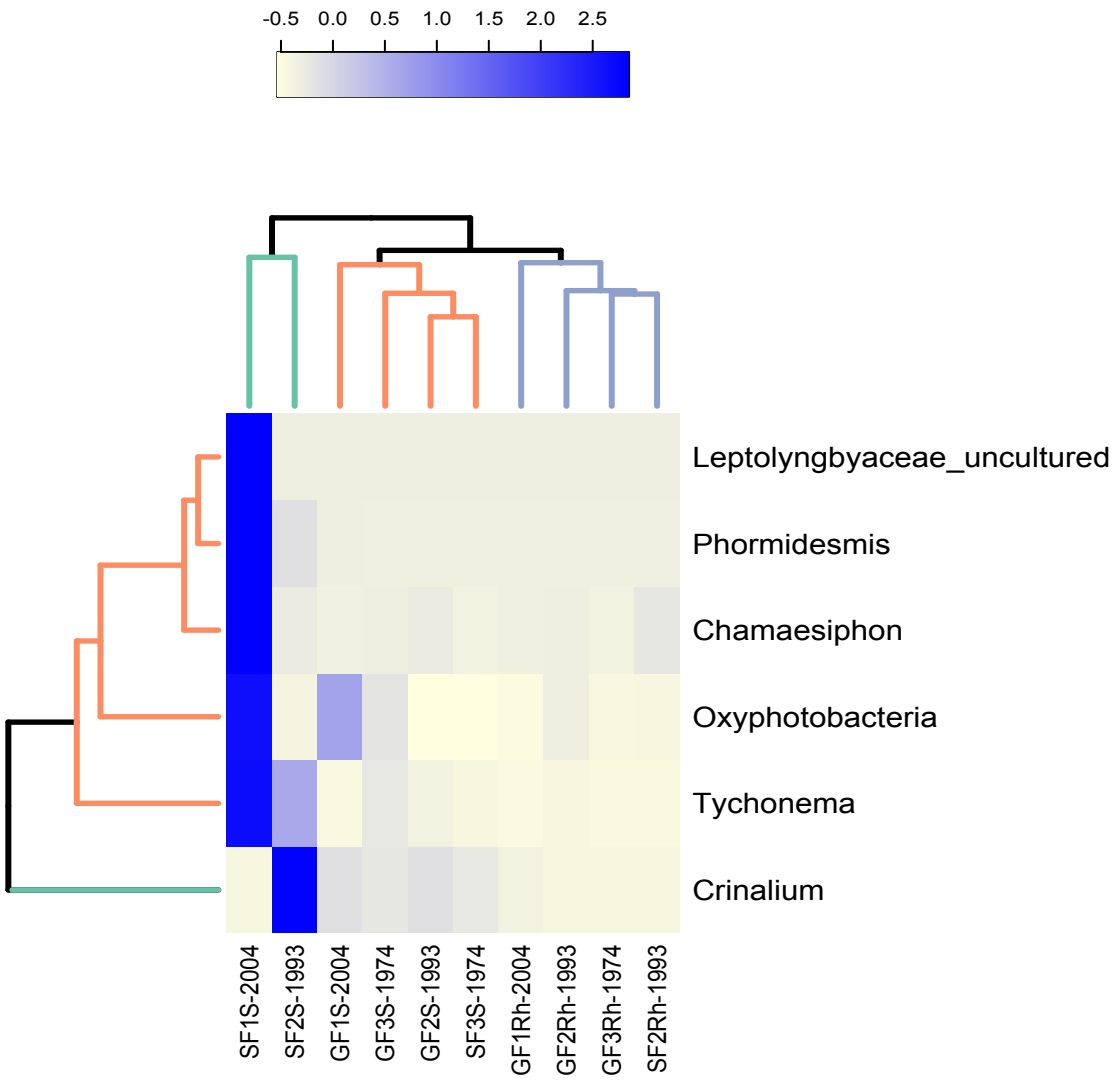


Figure 4.6: Hierarchical clustering of genera of *Cyanobacteria* in foreland samples. GF = glacier foreland, GFRh/Rhizo = Glacier foreland rhizosphere, SF = sideforland samples.

The significant difference in the distribution of the *Cyanobacteria* genera was also supported by the ANOSIM results between the sampling sites, upper (UL), lower (LL), glacier foreland (GF), side-foreland (SF) (*r = 0.308, p = 0.001;* **Figure 4.7**), between the sample types, cryoconite, rhizosphere and soil samples (*r = 0.584, p = 0.001;* **Figure 4.7**), between the ages along the foreland chrono-sequence (*r = 0.629, p = 0.001;* **Figure 4.7**), and between the altitude levels (*r = 0.255, p = 0.001;* **Figure 4.7**). The taxa diversity was highly affected by the ages along the foreland chrono-sequence, sample types, sampling

sites and altitude (*r = 0.629, r = 0.584, r = 0.308* and *r = 0.255;* **Figure 4.7**) respectively across the glacier and its foreland.

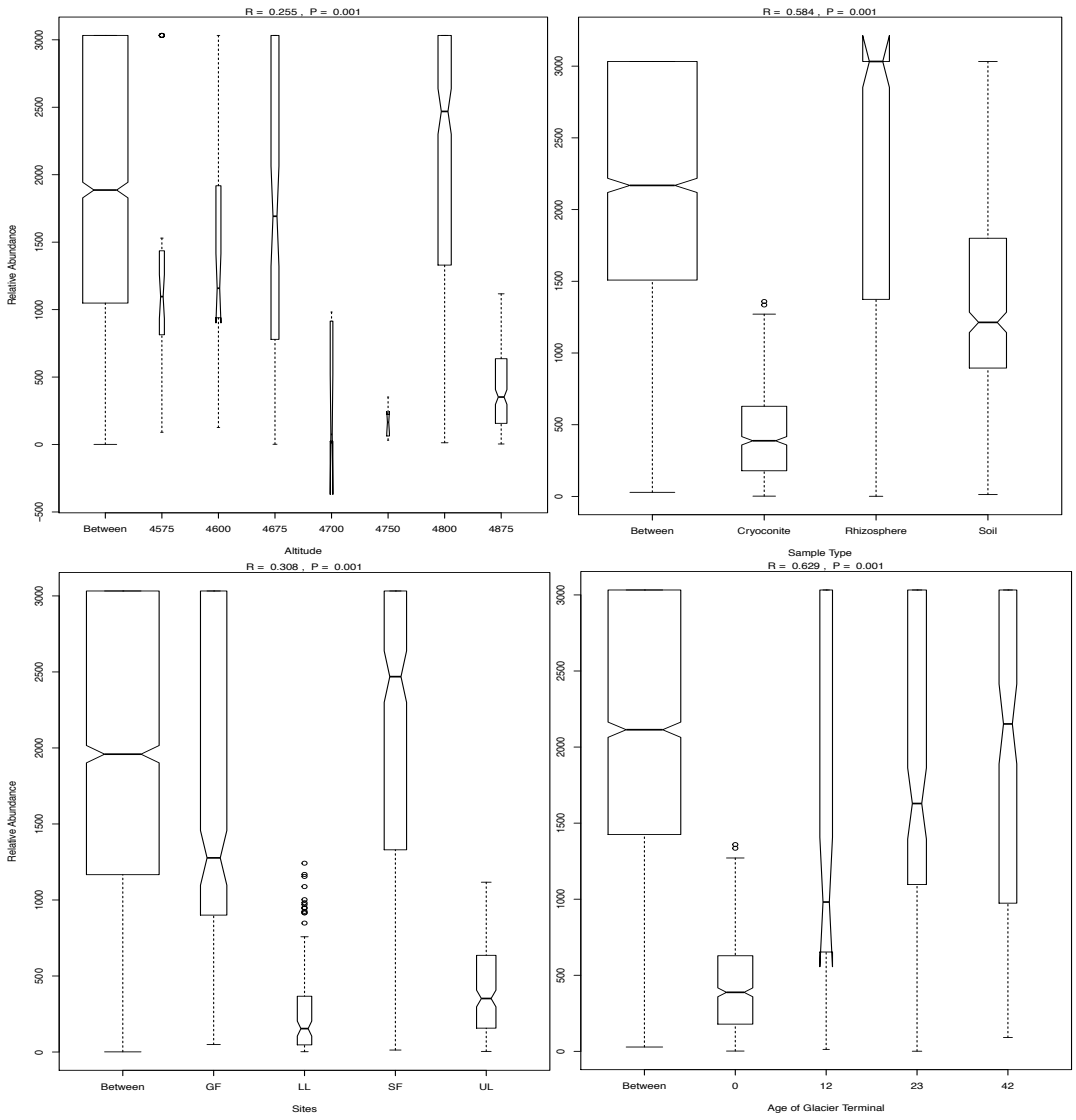


Figure 4.7: Differences in similarities in ASV composition of phylum *Cyanobacteria* between the sample types, altitude, sampling sites and age of the last glacier terminal based on the Analyses of Similarity (ANOSIM) method.

4.2.3 Abundance of *Proteobacteria* in the Glacier and its foreland

The analyses of the 16S rDNA of the glacier and foreland (soil and rhizosphere) samples revealed diverse genera of second most prominent phylum *Proteobacteria* (**Figure 4.8**). The lower site of the glacier (LG2) had high abundance of the *Proteobacteria* genera and the upper site had high community structures, that included the glacier specialists and generalized inhabitants like the vascular plant symbiont, genus *Mesorhizobium* that were highly abundant in the 12-year-old (soil and rhizosphere; 47 % and 18 %). Its abundant was, however, stable at the 23-year-old glacier foreland (GF2S) and 23-year-old side foreland (SF3S-2) soils (17.6 %) as in **Figure 4.8**.

Upper Lewis site had less diverse genera, but rich community structures ranging between 17 % to 30 % of genera *Rhizobacter* and *Rhodovastum* respectively. Genus *Rhodovastum* occurred in the glacier and the 12-year-old soil foreland sample with relative abundance of 30 % - 6.7 % respectively. Genus *Rhizobacter* was relatively high in the lower glacier site by 54 % to 17 % of the upper site. Its abundant, however, changed along the foreland soil gradient by 5 % and 14 % in the 12-year-old and 23-year-old soils respectively.

Genus *Acidiphilium* was highly abundant in the upper glacier site by 22 % to 20 % in the lower glacier site. Its abundant in the foreland sample changed by relative abundance of 3 %, 11 % and 12 % in the 12-year-old soil, 23-year-old soil and 42-year-old soil respectively. Genus *Acidiphilium* only occurred in the 42-year-old rhizosphere (SF3R) sample by a low relative abundance of 3 %. Genus *Polymorphobacter* only occurred in the S4LG1 sample, which also showed high community structure in the upper glacier site (**Figure 4.8**). Other genera such as *Sphingomonas* were only present in the glacier foreland and side foreland samples. The foreland samples had diverse genera with lower community structure while the glacier samples had a few genera with high community structures.

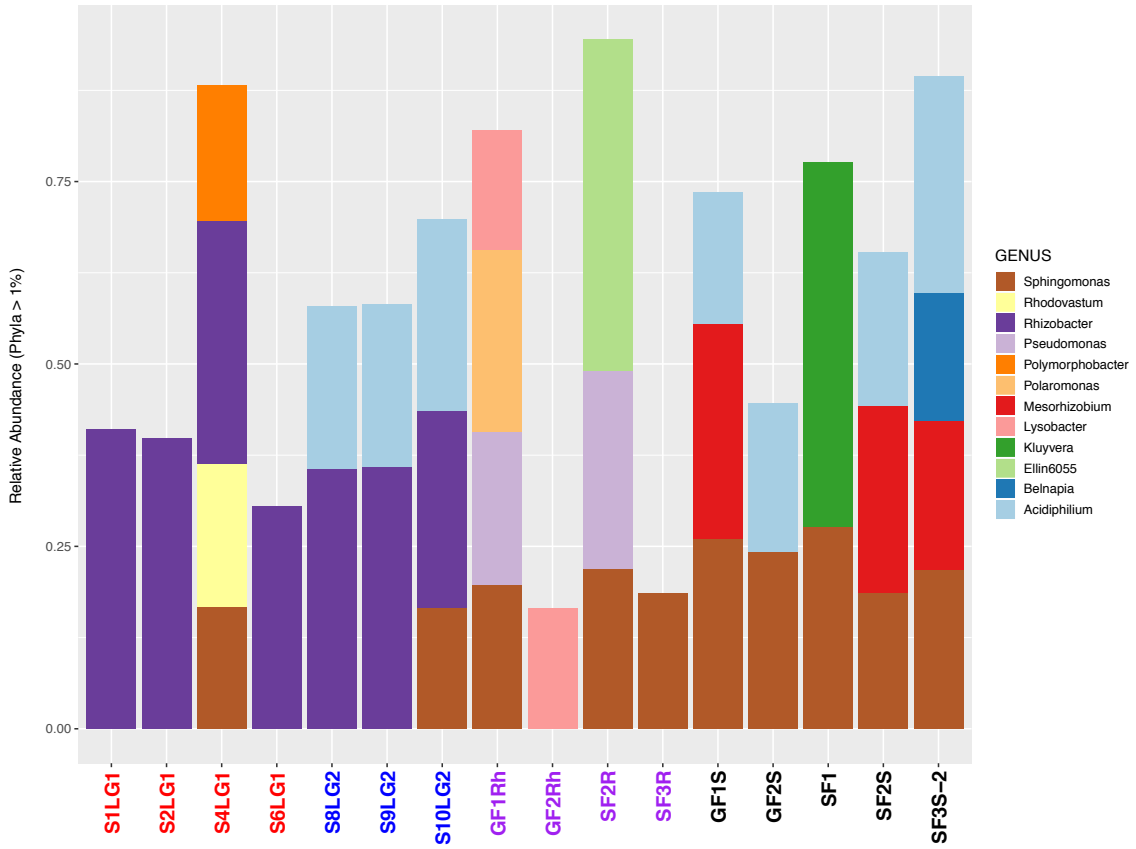


Figure 4.8: Genera of phylum *Proteobacteria* from the glacier and foreland samples. LG1 = Upper Lewis site, LG2 = Lower Lewis site, GF = glacier foreland, Rh = rhizosphere, SF = Side foreland.

Class *Gammaproteobacteria* was prominent in the glacier upper site, class *Alphaproteobacteria* dominated the foreland samples while class *Deltaproteobacteria* was presented with low abundance across the glacier and foreland samples (**Figure 4.9**).

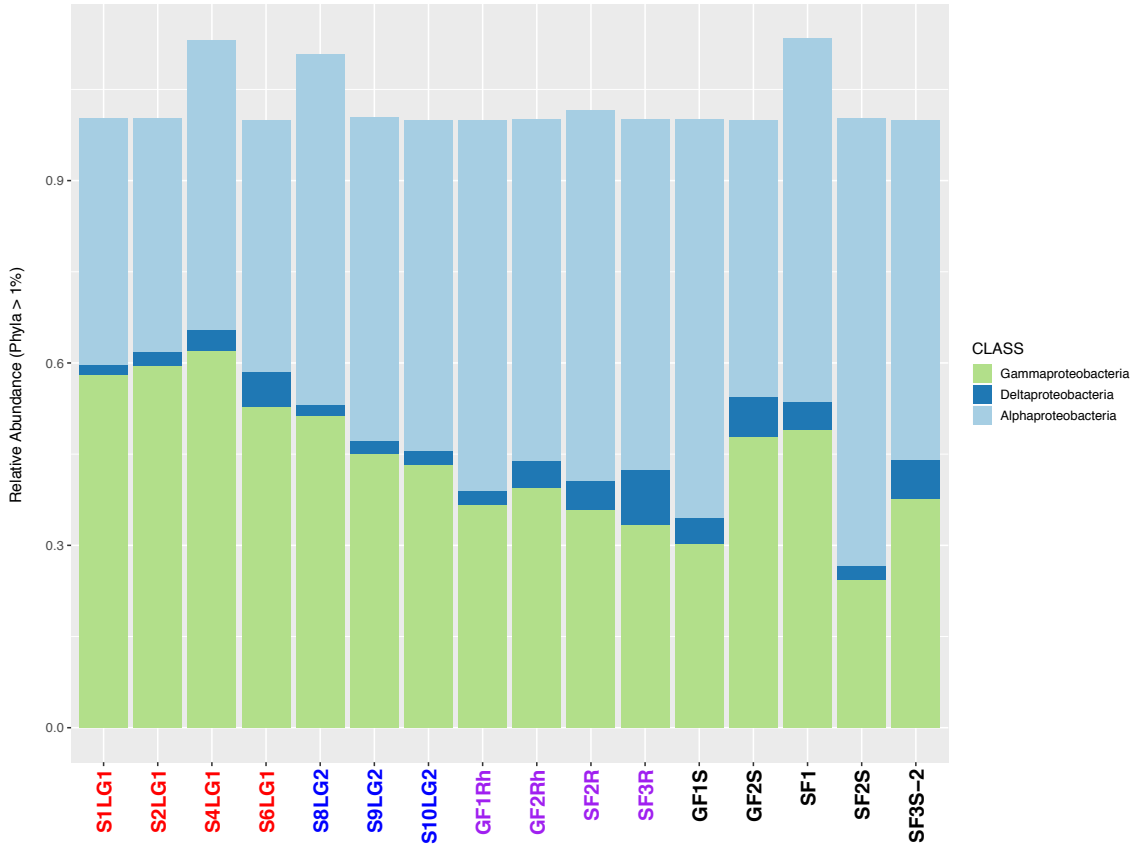


Figure 4.9: Predominant *Proteobacteria* class on the glacier and foreland samples. LG = Lewis Glacier, GF = glacier foreland, SF = side foreland, Rh = Rhizosphere, S = Soil.

Hierarchically, samples were clustered according to the sites of sample and class distribution. Upper glacier sites were clustered together with the 23 and 42-year-old foreland samples while the 23-year-old side-foreland soil sample (SF2S_1993) clustered independently as well as the class *Deltaproteobacteria* that was presented in low abundance (**Figure 4.10**). The 23-year-old side-foreland soil sample had low class diversity with high abundance of class *Alphaproteobacteria.*

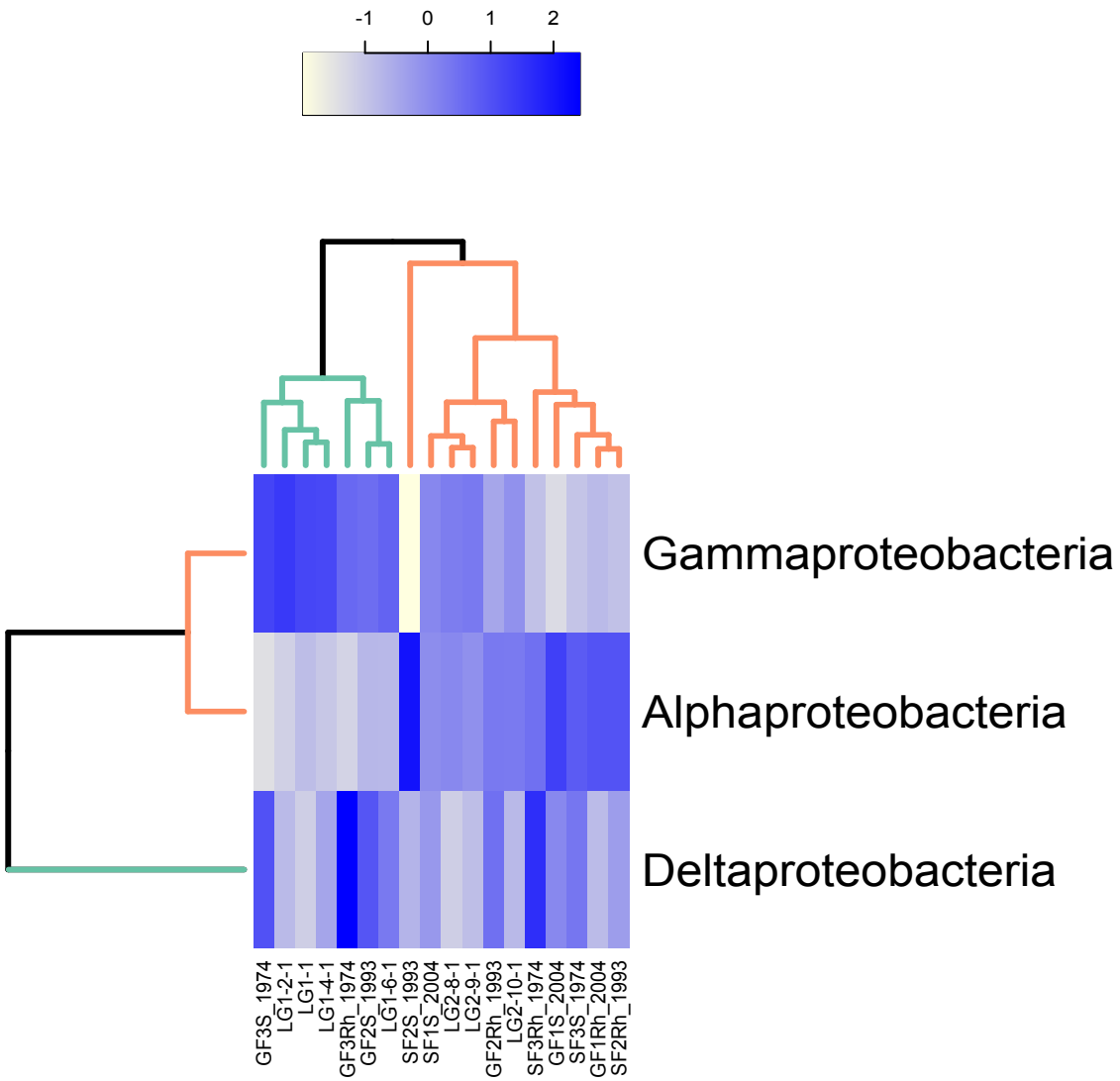


Figure 4.10: Hierarchical clustering of the class of *Proteobacteria* on the glacier and foreland samples. GF = glacier foreland, SF = side foreland, Rh = Rhizosphere, S = Soil.

4.2.3.1 Composition and Abundance of *Alphaproteobacteria* in the glacier and its foreland

Though, *Alphaproteobacteria* were the most abundant class in the phylum *Proteobacteria,* the hierarchical clustering of its 16S rDNA clustered the major genera

according to the sampling site and age for the foreland samples (**Figure 4.11 a** and **b**). Upper glacier sites were clustered together, except LG1.6.1 which clustered together with the lower glacier samples (**Figure 4.11a**). Genera *Mesorhizobium* and *Sphingomonadaceae* were hierarchically unique to sample LG1.4.1 which was clustered separately from the other clusters. Sample LG1.4.1 was also unique due high diversity of genera and the microbial community structure (**Figure 4.11a**).

In the foreland, soil samples were also clustered together, except the 23-year-old and 42-year-old glacier foreland soil samples that were clustered together with the rhizosphere samples (**Figure 4.11b**). Genera *Polymorphobacter* and *Rhodovastum* were unique to 12-year-old side-foreland sample and were clustered together with SF1S_2004 as an outstanding sample (**Figure 4.11b**).

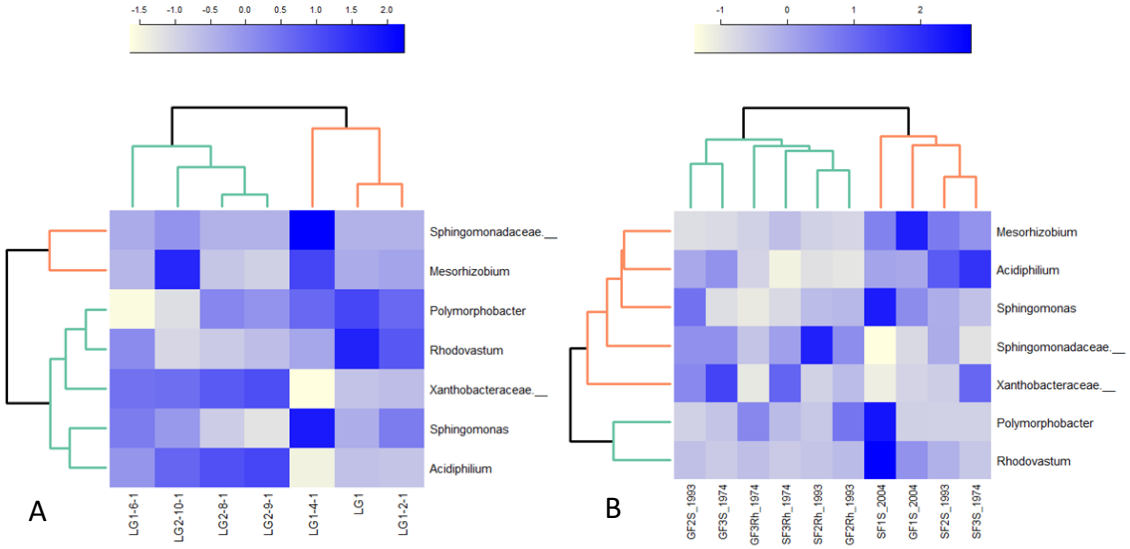


Figure 4.11: Hierarchical clustering of the genera of class *Alphaproteobacteria* in the glacier samples (A) and in the foreland samples (B). LG = Lewis Glacier, GF = glacier foreland, SF = side foreland, Rh = Rhizosphere, S = Soil.

4.2.3.2 Composition and Abundance of *Deltaproteobacteria* in the glacier and its foreland

Class *Deltaproteobacteria* was presented in the samples with low abundance across the glacier and foreland samples. Hierarchical clustering based on the Bray-Curtis

dissimilarity index, revealed this class to be distributed almost evenly and independently to each of the glacier and foreland samples (**Figure 4.12a** and **b)**. Genera *Pajaroellobacter* and *Haliangum* were, however, clustered together and were lowly presented in samples LG1.4.1, which was also clustered separately (**Figure 4.12a**).

In the foreland, genus *Geobacter* was independenly clustered from other genera as it was relatively abundant in samples GF3Rh_1974 and SF2S_1993 (**Figure 4.12b**). The foreland samples were also clustered according to site and age across the foreland chrono-sequence. The 12-year-old sample SF1S_2004 had low ASV diversity and was clustered separately from the other samples.

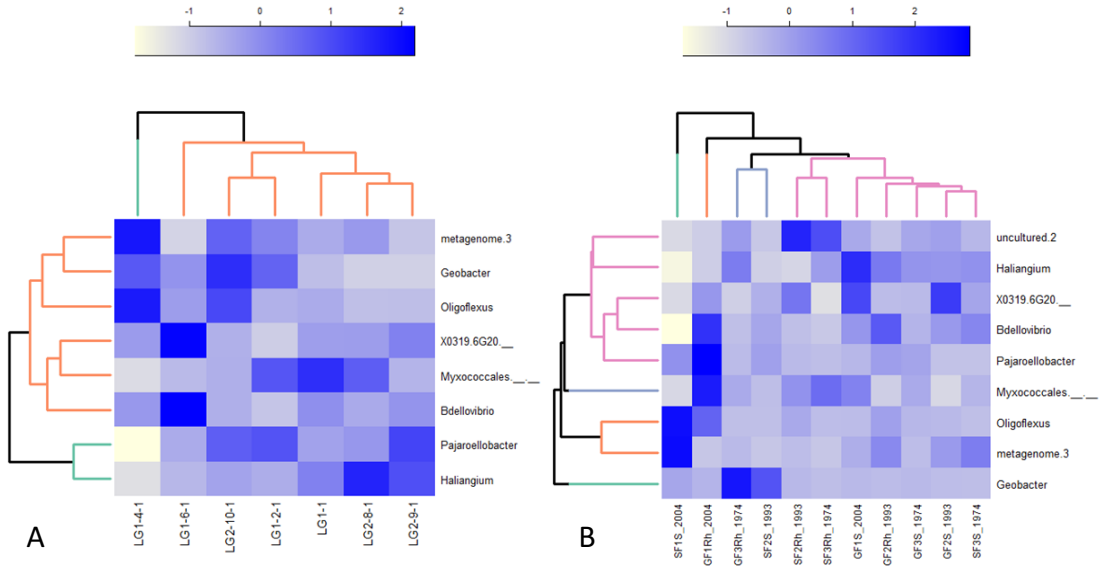


Figure 4.12: Hierarchical clustering of the major genera of class *Deltaproteobacteria* in the glacier samples (A) and in the foreland samples (B). LG = Lewis Glacier, GF = glacier foreland, SF = side foreland, Rh = Rhizosphere, S = Soil.

4.2.3.3 Composition and Abundance of *Gammaproteobacteria* in the glacier and its foreland

Class *Gammaproteobacteria* was the second most prominent class in the phylum *Proteobacteria* after *Alphaproteobacteria.* Hierarchically, most of the genera are evenly distributed across the glacier sample sites. However, afew of the genera are clustered

independently of the others. Genus *Rhizobacter* is more abundant in the lower glacier sites and relatively low in the upper glacier sites (**Figure 4.13a**). Sample LG2.10.1 is clustered separately and it has low ASV diversity with high community structure of genera, *Methylotenera, Noviherbaspirillum* and genus to *Burkholderiaceae* (**Figure 4.13a**). the foreland cluster of the class *Gammaproteobacteria* is, however, scattered, indicating low abundance of the class in the foreland chron-sequence (**Figure 4.13b**). Unlike in the glacier sample, the clusters of the ASVs have slightly interchanged. The two major clusters include, *Polaromonas, Rhizobacter* and *Ellin6067* are clustered together while genera *Methylotenera, Noviherbaspirillum* and genus to *Burkholderiaceae* were also clustered together. Some of the foreland samples, SF1S_2004, also showed low ASV diversity, though high community structure of specific genus (**Figure 4.13b**).

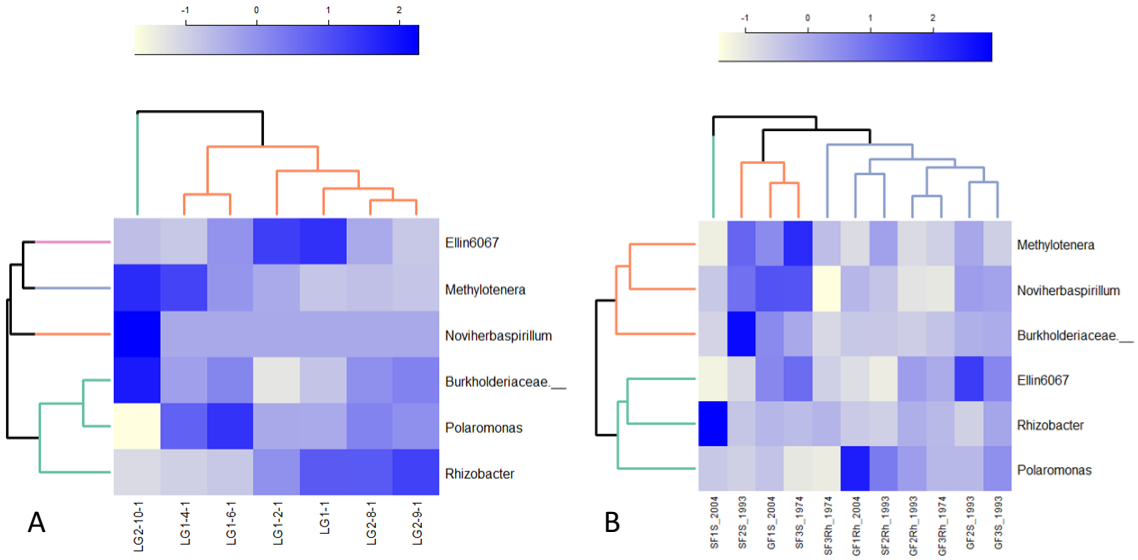


Figure 4.13: Hierarchical clustering of the major genera of class *Gammaproteobacteria* in the glacier samples (A) and in the foreland samples (B). LG = Lewis Glacier, GF = glacier foreland, SF = side foreland, Rh = Rhizosphere, S = Soil.

The significant difference in the composition and distribution of the phylum *Proteobacteria,* its classes and major genera was also supported by the ANOSIM results between the between the altitude levels (*r = 0.389, p = 0.001;* **Figure 4.14**), sample types,

cryoconite, rhizosphere and soil samples (*r = 0.73, p = 0.001;* **Figure 4.14** sampling sites, upper (UL), lower (LL), glacier foreland (GF), side-foreland (SF) (*r = 0.469, p = 0.001;* **Figure 4.14**), and between the ages along the foreland chrono-sequence (*r = 0.681, p = 0.001*; **Figure 4.14).** The taxa diversity was highly affected by the sample type, ages along the foreland chrono-sequence, sampling sites and altitude (*r = 0.73, r = 0.681, r = 0.469* and *r = 0.389;* **Figure 4.14**) respectively across the glacier and its foreland.

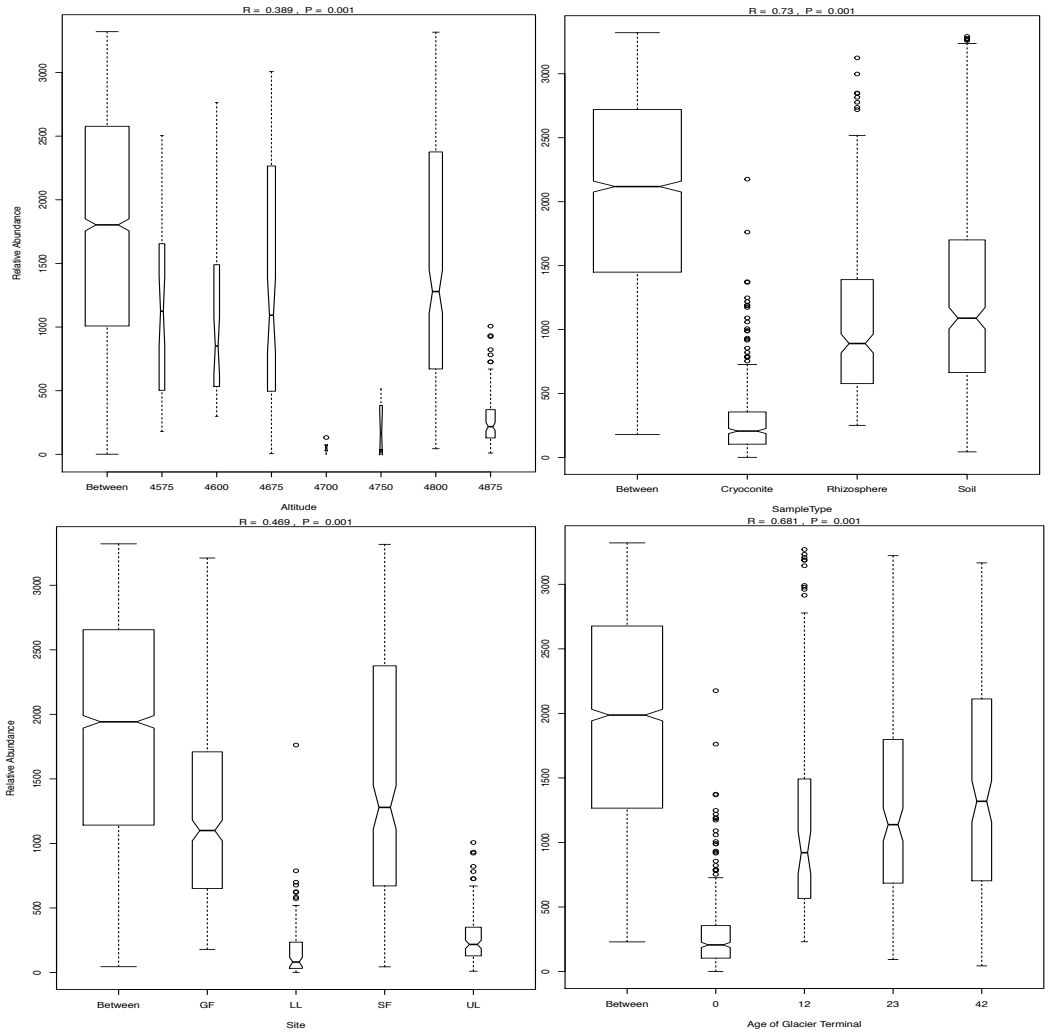


Figure 4.14: Differences in similarities in ASV composition of phylum *Proteobacteria* between the sample types, altitude, sampling sites and age of the last glacier terminal based on the Analyses of Similarity (ANOSIM) method.

4.2.4. Abundance of *Bacteroidetes* in the Glacier

ASVs that aligned to phylum *Bacteroidetes* were filtered and the five genera had a relative abundance of above average. The five genera were then picked for hierarchical cluster analyses based on Bray-Curtis dissimilarity analyses. The heat map showed the genera to be sparsely distributed across the glacier samples (**Figure 4.15**). Only genus *Spirosoma* and *Fibrella* were clustered together as they appeared to dominate the lower glacier samples. The other genera were clustered independently across the upper and the lower glacier sites.

In the foreland, 15 known genera were hierarchically clustered (**Figure 4.16**). Most of the genera were also clustered together except a few that were clustered separately from the common clade. Some samples were also out-grouped. Sample SF1-5 had invisible ASVs and therefore could only be clustered separately. Generally, the major genera were evenly across the foreland chrono-sequence (**Figure 4.16**).

The significant difference in the composition and distribution of the phylum *Bacteroidetes* was supported by the ANOSIM results between the between the altitude levels ($r = 0.446$, $p = 0.001$; **Figure 4.17**), sample types, cryoconite, rhizosphere and soil samples ($r = 0.683$, $p = 0.001$; **Figure 4.17**), sampling sites, upper (UL), lower (LL), glacier foreland (GF), side-foreland (SF) ($r = 0.585$, $p = 0.001$; **Figure 4.17**), and between the ages along the foreland chrono-sequence ($r = 0.631$, $p = 0.001$; **Figure 4.17**). The taxa composition and diversity were highly affected by the sample type, ages along the foreland chrono-sequence, sampling sites and altitude ($r = 0.683$, $r = 0.631$, $r = 0.585$ and $r = 0.446$) respectively across the glacier and its foreland (**Figure 4.17**).

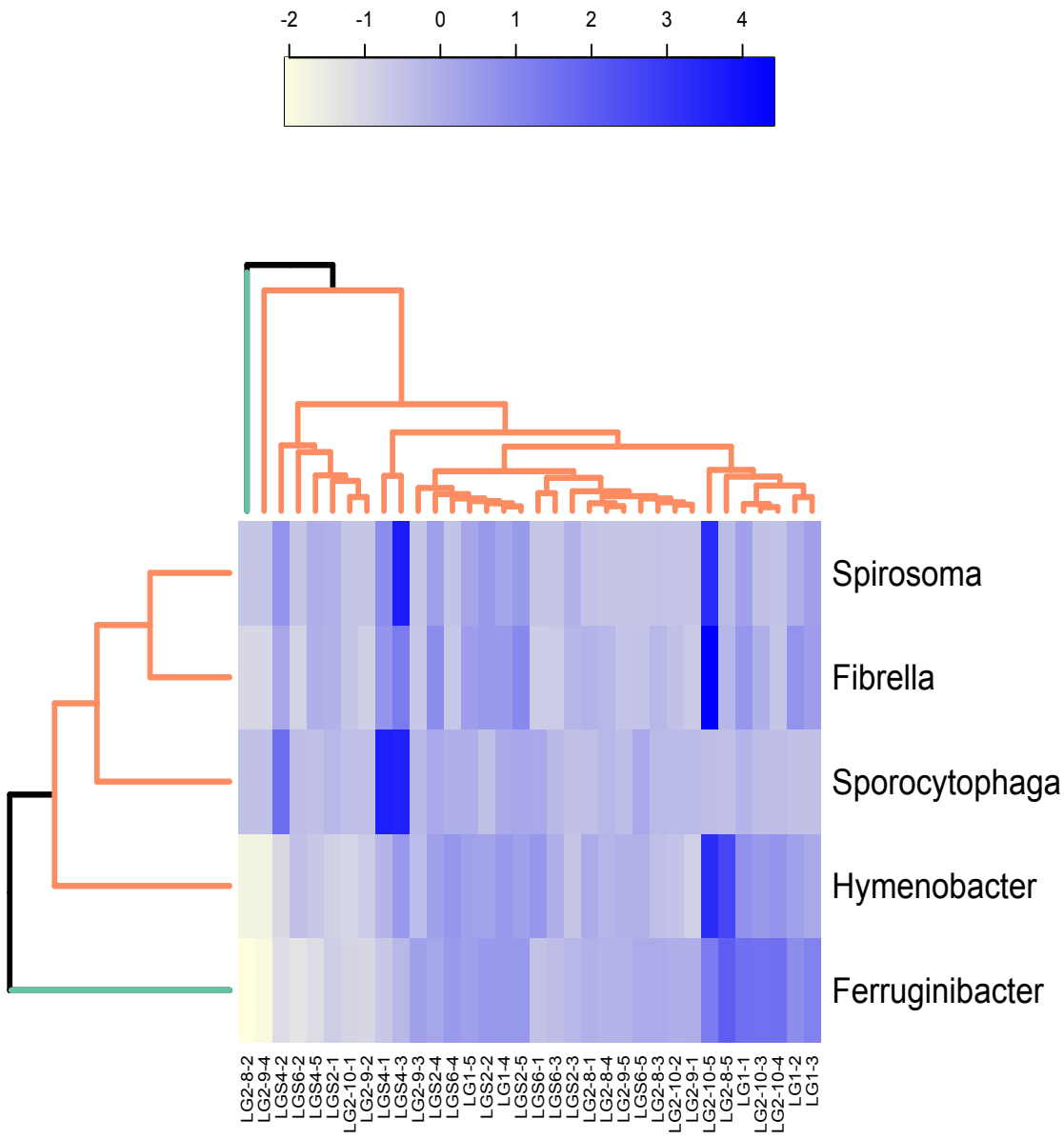


Figure 4.15: Hierarchical clustering of 16S rDNA of major genera of the Phylum *Bacteriodetes* on glacier samples. LG = Lewis glacier.

4.2.5. Abundance of *Bacteriodetes* in the Glacier foreland

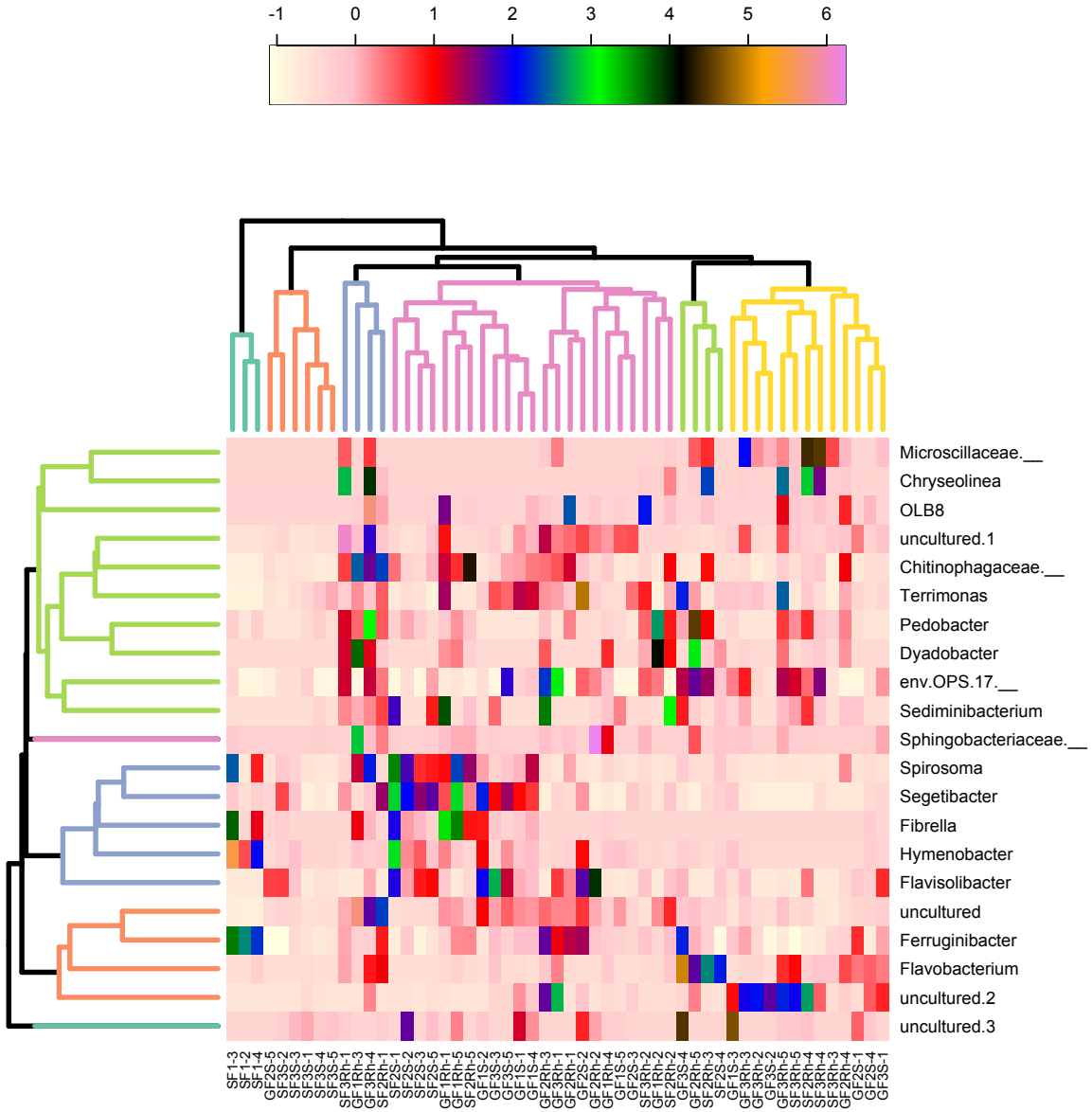


Figure 4.16: Hierarchical clustering of the major genera of the Phylum *Bacteroidetes* on foreland samples. GF = glacier foreland, SF = side foreland, Rh = Rhizosphere, S = Soil.

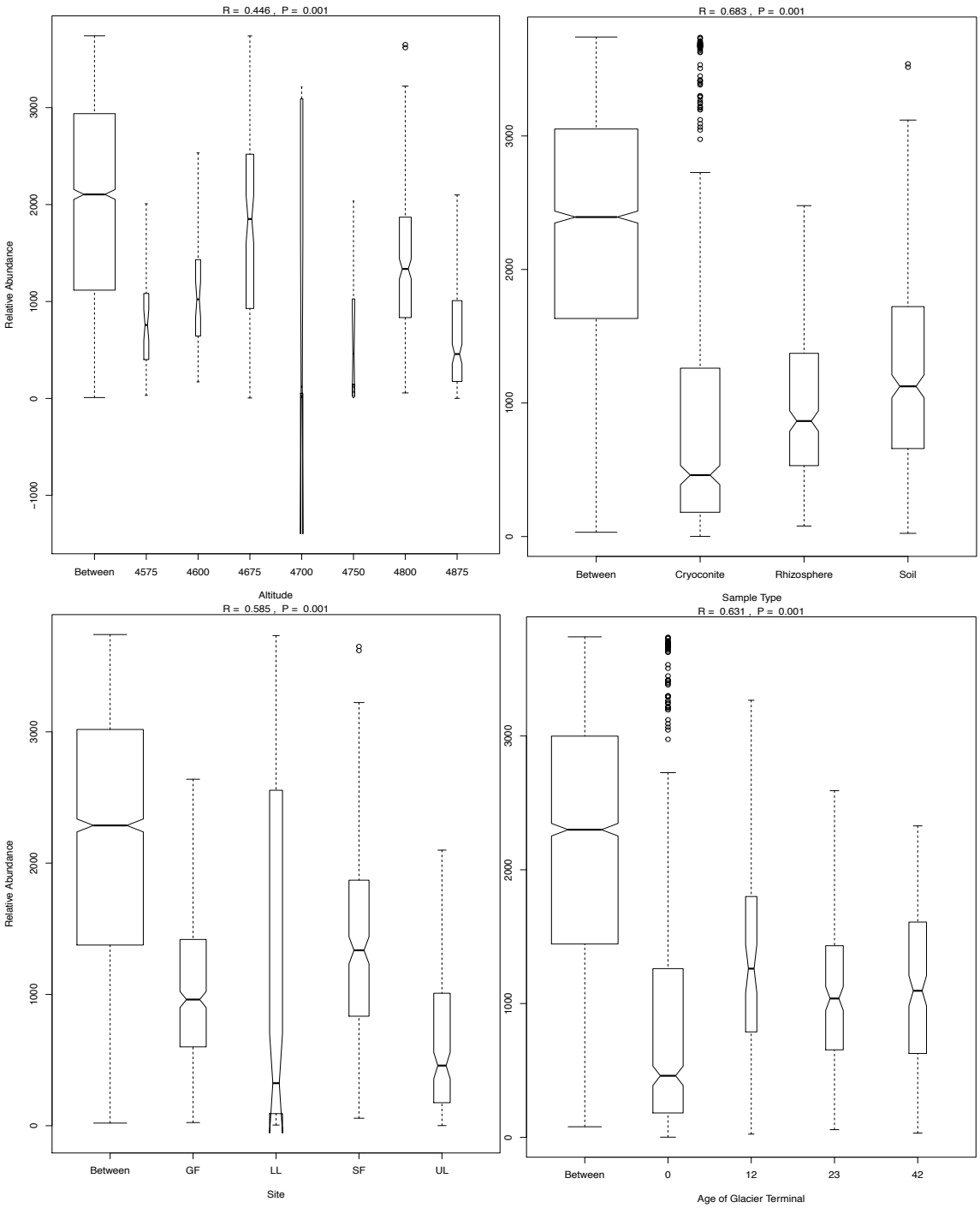


Figure 4.17: Differences in similarities in ASV composition of phylum *Bacteroidetes* between the sample types, altitude, sampling sites and age of the last glacier terminal based on the Analyses of Similarity (ANOSIM) method.

4.2.6. Abundance of *Actinobacteria* in the Glacier and Glacier foreland

Eleven genera of *Actinobacteria* were revealed from the 16S rDNA analyses of the glacier and foreland (soil and rhizosphere) samples. Generally, the distribution pattern of the genera of *Actinobacteria* was different from the phylum *Proteobacteria.* The glacier sites demonstrated low diversity of the major genera, but high number of the community structure of the clustered genera (3 % - 41 %) as in **Figure 4.18.**

Parafrigoribacterium is the prominent genus in the glacier with a relative abundance of 25 % in lower site and 41 % in the upper site. Genus *Oryzihumus* is also prominent in the glacier site (15 %; lower site, 17 % upper site), but poorly distributed in the foreland sites (2 % - 4 %).

Hierarchically, glacier samples formed two major clusters. The clusters majorly consist of the *Rhodococcus, Nakamurella,* and *Oryzihumus.* Sample LG2-9-4 had low abundance of ASVs and was out grouped from other samples (**Figure 4.19**). The foreland samples were also clustered together in a heatmap showing diversified distribution of *Actinobacteria* genera across the foreland age chrono-sequence. The samples were clustered according to the sample type. Rhizosphere samples were clustered together while soil samples also formed a clade **(Figure 4.20).** The heatmap showed that, the 12-year-old soil sample had high density of community structures as compared to the 42-year-old soil samples (**Figure 4.20**). SF1-5 formed an out group as it shared some of the glacier endogenous and exogenous communities.

The significant difference in the composition and distribution of the phylum *Actinobacteria* was supported by the ANOSIM results between the between the altitude levels (*r = 0.446, p = 0.001;* **Figure 4.21**), sample types, cryoconite, rhizosphere and soil samples (*r = 0.722, p = 0.001;* **Figure 4.21**), sampling sites, upper (UL), lower (LL), glacier foreland (GF), side-foreland (SF) (*r = 0.509, p = 0.001;* **Figure 4.21**), and between the ages along the foreland chrono-sequence (*r = 0.682, p = 0.001;* **Figure 4.21**).

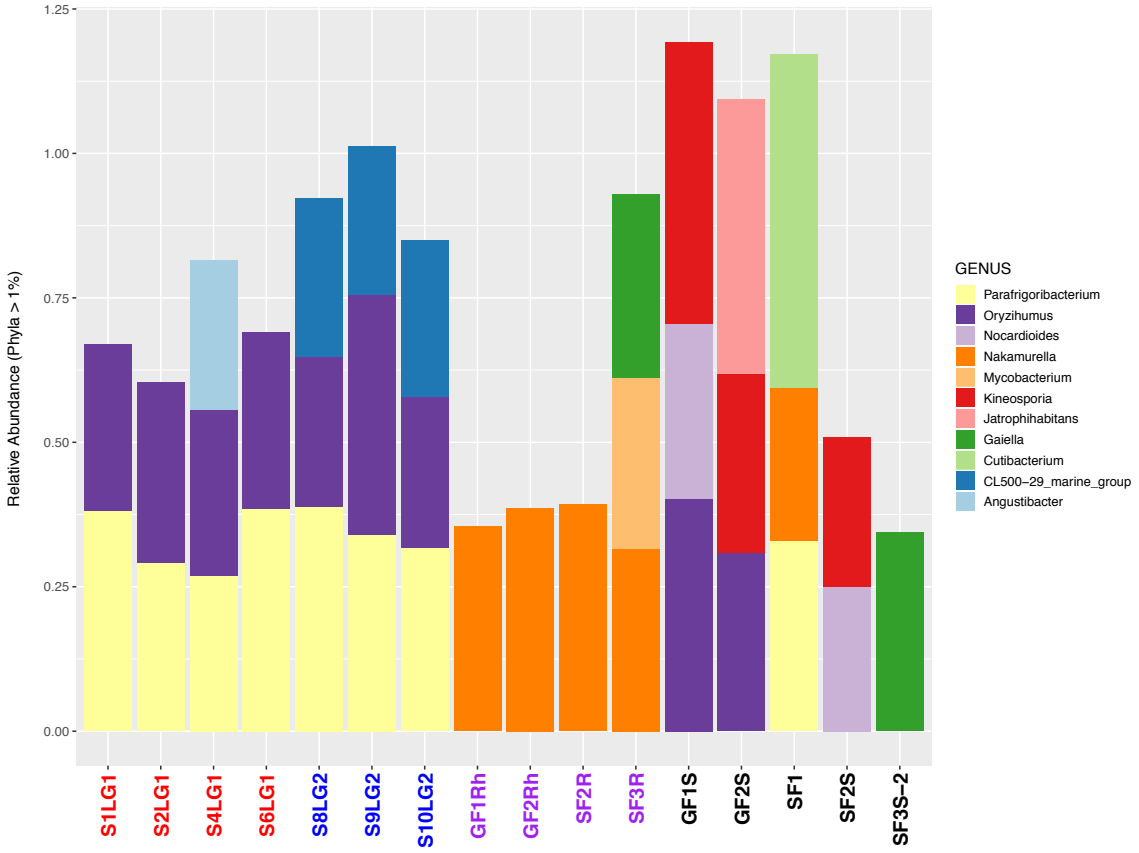


Figure 4.18: Genera of the phylum *Actinobacteria* from the glacier and foreland samples. LG1 = Upper Lewis site, LG2 = Lower Lewis site, GF = glacier foreland, Rh = rhizosphere, SF = Side foreland.

The taxa composition and diversity were highly affected by the sample type, ages along the foreland chrono-sequence, sampling sites and altitude (*r = 0.683, r = 0.631, r = 0.585* and *r = 0.446*) respectively across the glacier and its foreland (**Figure 4.21**).

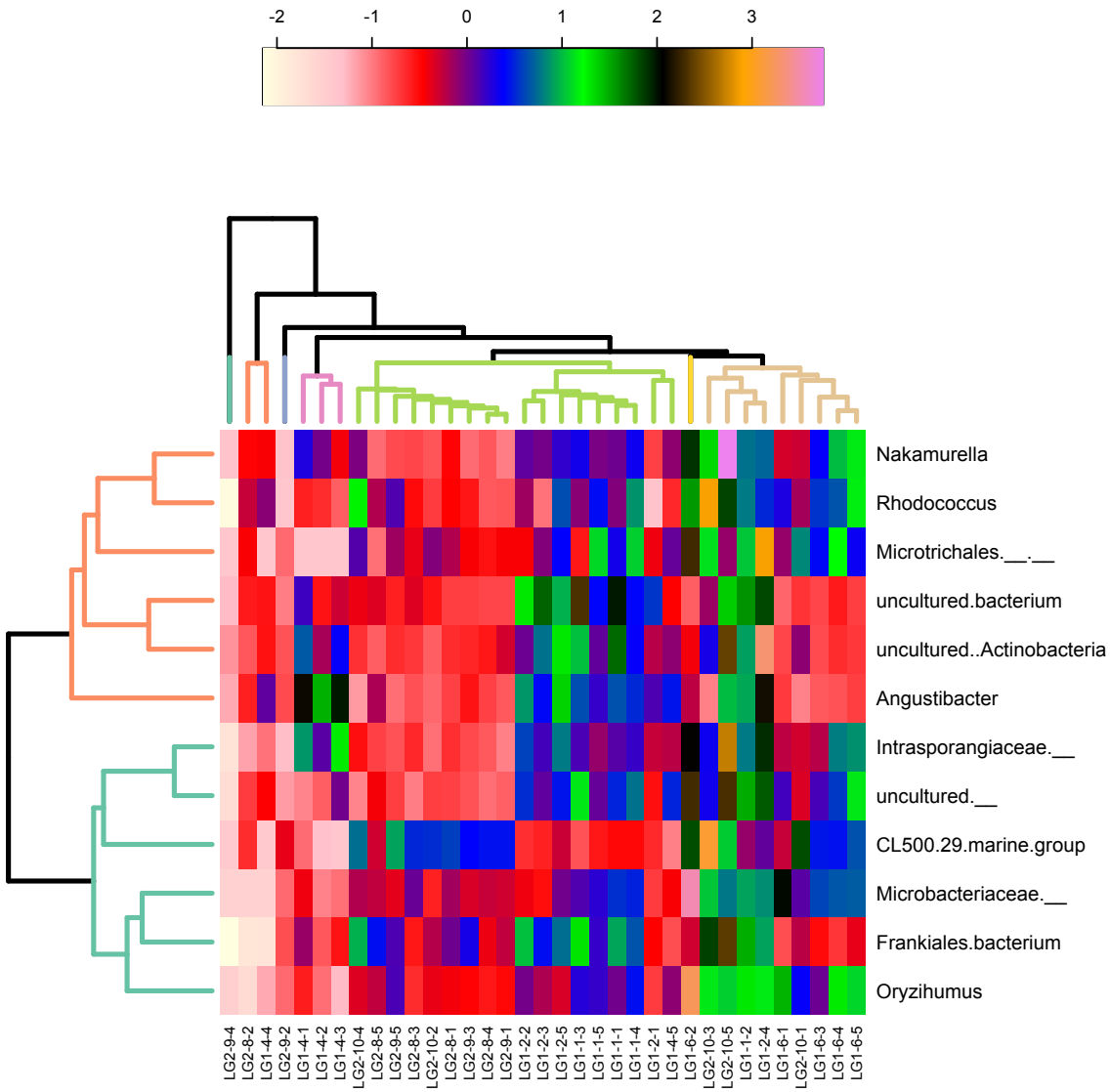


Figure 4.19: Hierarchical clustering of the major genera of the Phylum *Actinobacteria* on glacier samples. LG = Lewis glacier.

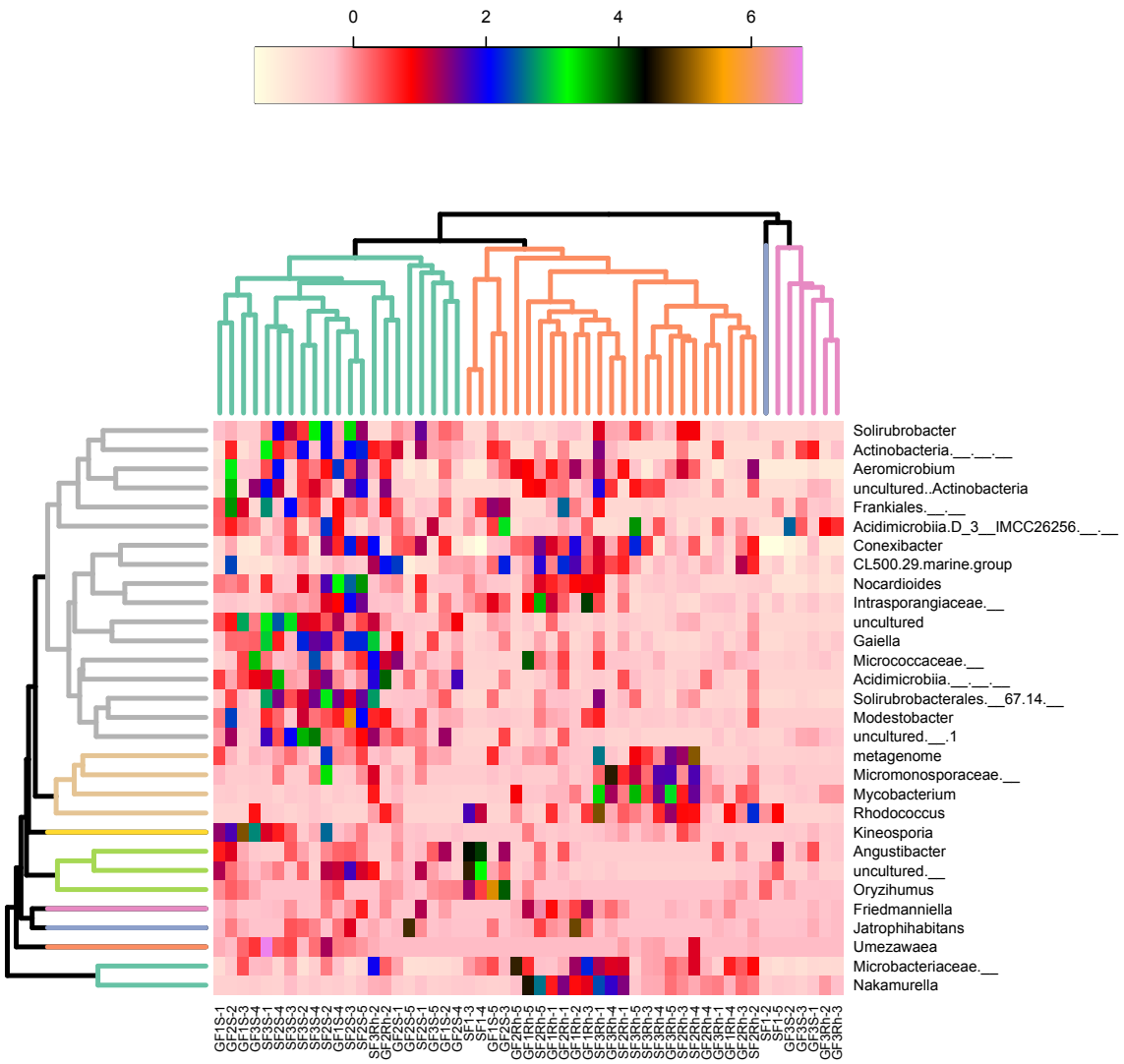


Figure 4.20: Hierarchical clustering of the major genera of the Phylum *Actinobacteria* on foreland samples. GF = glacier foreland, SF = side foreland, Rh = Rhizosphere, S = Soil.

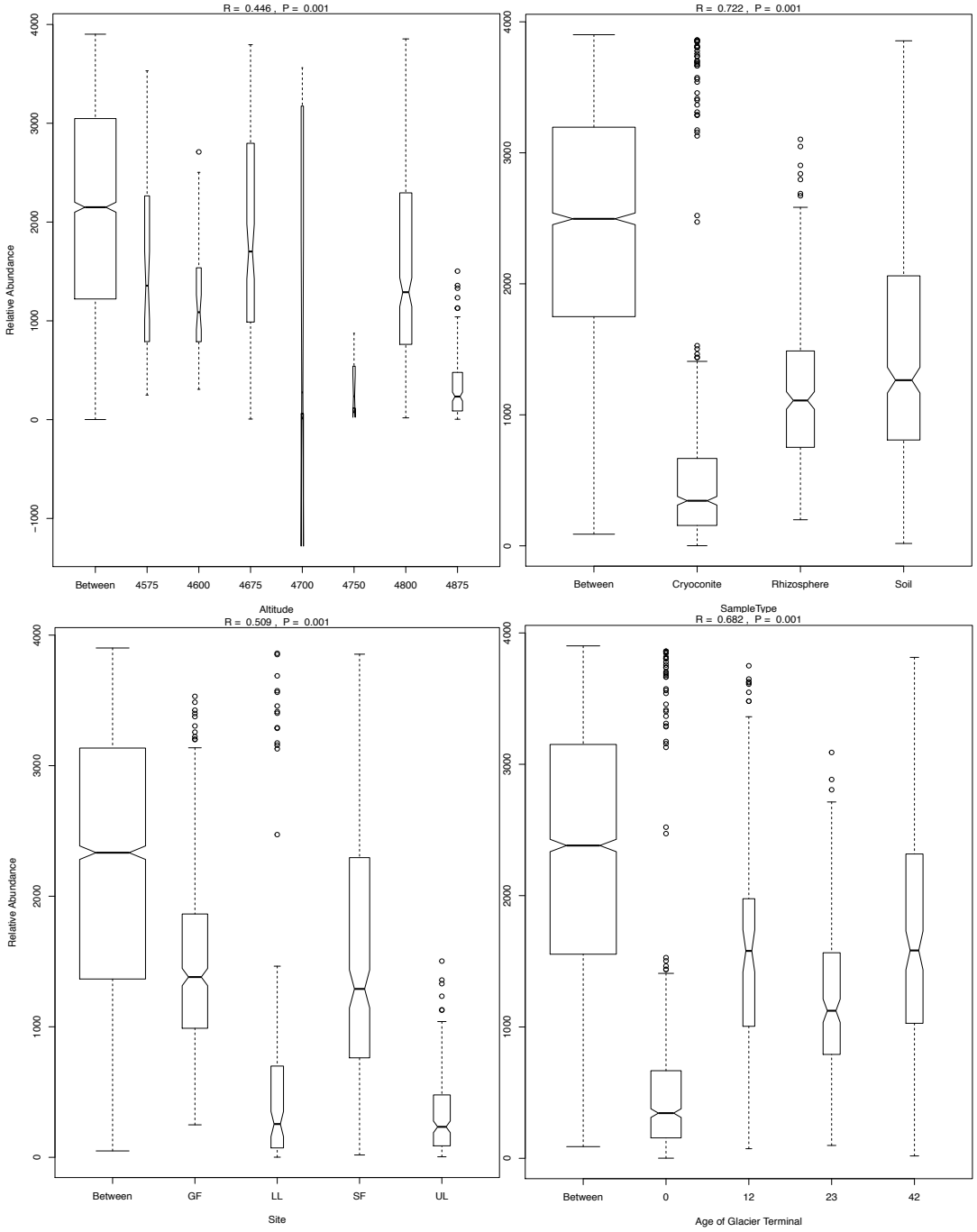


Figure 4.21: Differences in similarities in ASV composition of phylum *Actinobacteria* between the sample types, altitude, sampling sites and age of the last glacier terminal based on the Analyses of Similarity (ANOSIM) method.

4.3. Diversity of prokaryotes in the supraglacial snow on Lewis Glacier

4.3.1 Microscopy of the snow and ice samples

The shapes and pattern of the filtrates that were observed revealed known features of flagellated *Cyanobacteria,* a parasite tardigrade, red algae and mesh-work of organisms (**Plate 4.1a, b, c and d**) respectively. When the 200 *µl* melt ice and dissolved soil were inoculated into specific media for physiological analyses, growth was observed to be rapid at 25°C, at 10°C the growth was relatively gradual while at 4°C which simulated the glacier conditions, growth was absolutely slow, but steady.

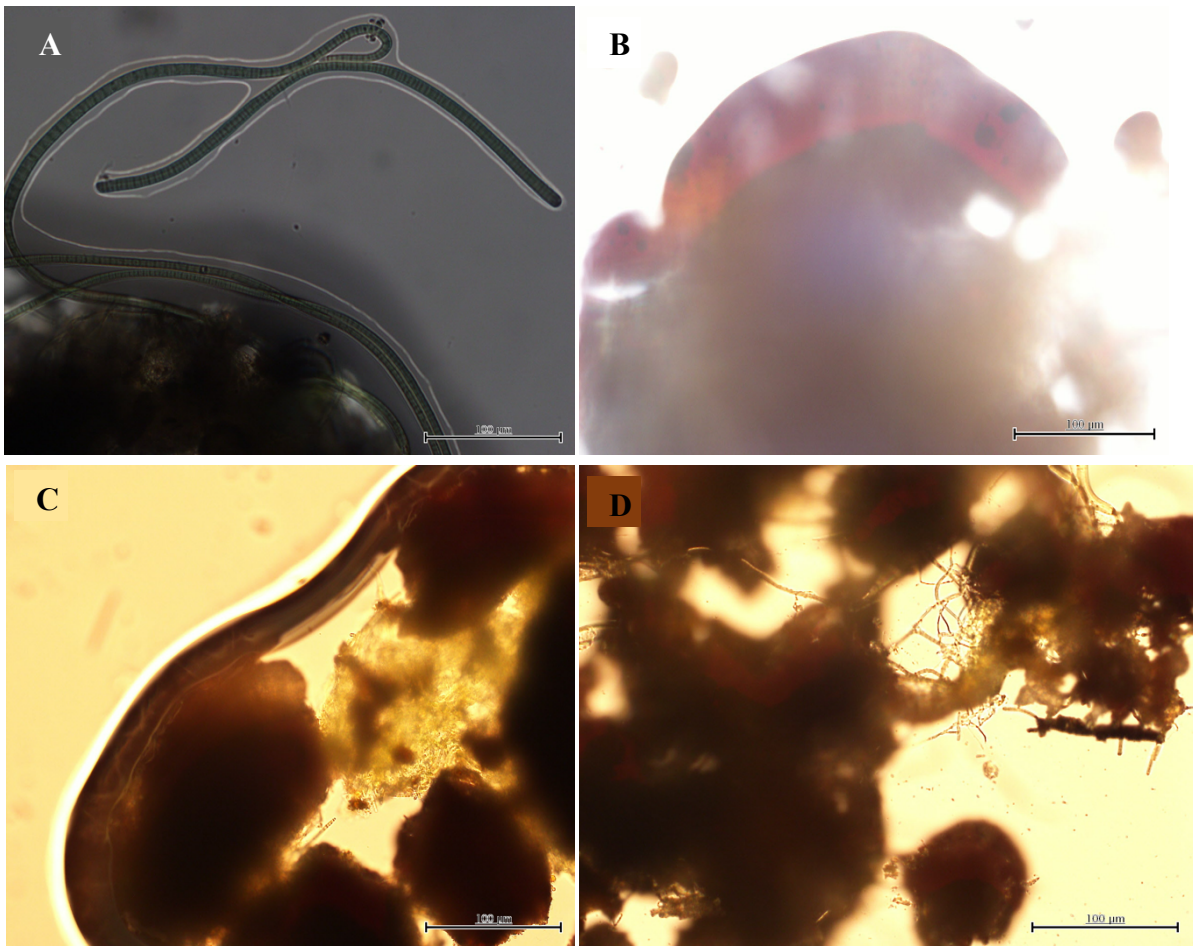


Plate 4.1: Glacier microscopy showing *Cyanobacteria* (A), tardigrades (B), red algae (C) and mesh-work of organisms (D).

The microscopy of the glacier terminal melt water and soil samples revealed the availability of fungal spore, short flagellate, *Cyanobacteum* species and thin long organism (**Plate 4. 2a, b, c and d).** The *Cyanobacteum* species was sampled from the primary glacier deposits, the short flagellate was sampled from the bare soil and the thin long organism was sampled from the plant rhizosphere.

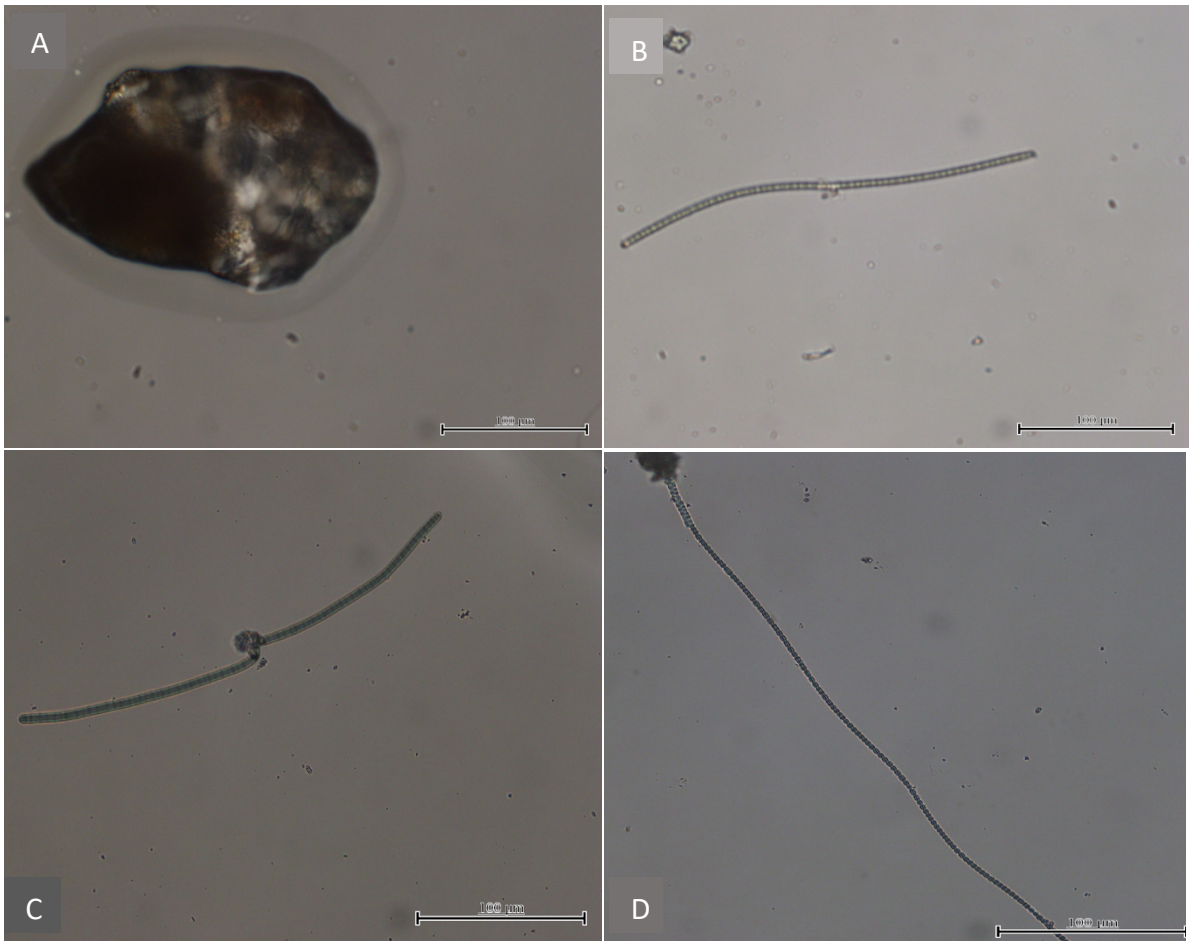


Plate 4.2: Melt water microscopy showing fungal spore (A), short flagellates (B), thin green *Cyanobacteria* (C) and thin long flagellated organism (D).

4.3.2 Molecular analyses of the isolates

The partial analyses of 16S rDNA revealed a total of seventeen isolates from snowpack of Lewis glacier. The isolates were phylogenetically diverse and affiliated to known members of the phyla *Firmicutes, Proteobacteria,* and *Actinobacteria* (**Table 4. 1** and

Figure 4.22). Sequence similarity analyses of the 16S rDNA against NCBI Genbank database via Blastn showed similarities were more than 96% (**Table 4. 1**).

More than 50% of the isolates were affiliated genus *Bacillus* (> 96 %) sequence identity. Isolates, Lewis_bac_5, 6, 13 and 21 had 100% sequence similarity to genus *Stenotrophomonas* and they formed 23.5% of the total isolates. Isolates, Lewis_bac_2, 10, 19 and 20 were 100% similar to genera *Cryobacterium, Paenibacillus, Subtercola* and *Arthrobacter* respectively. Other isolates, Lewis_bac_2, Lewis_bac_19 and Lewis_bac_20 were 100% to genera *Cryobacterium, Subtercola/Agreia* and *Arthrobacter* respectively, which are glacier and polar zone colonies. They formed 17.6% of the total number of 17 isolates. Another 76.5% of the isolates had more than 96% similarity percentage to terrestrial and aquatic species while the other. The remaining 5.9% of the isolates were100% similar to genus *Stenotrophomonas*, which is a human gastrointestinal tract colonizer.

The tree was rooted with *Methanoculleus thermophilus* (AB065297) as an outgroup from the glacier isolates (**Figure 4.22**). The phylogenetic tree formed four sub-groups clustered under two major clusters. The clusters were described by the most possible sources of the isolates that were compared against the NCBI database. Isolates which were closely related to the snow inhabitants were clustered together in a single sub-cluster with a bootstrap value of 100%. Another cluster captured the isolates that were closely related to known members of the terrestrial and aquatic ecosystems. They were also clustered together in sub-groups describing species and strains of various genera (**Figure 4.22**). The phylo-tre clustered Lewis_bac_5, 6, 13 and 21 together with the genus *Stenotrophomonas,* which is a human gut pathogen.

Table 4.1: Taxonomic affiliation and percentage sequence similarities of bacterial isolates with closest relatives from the Genbank database.

Sample ID	Accession No.	Closest taxonomic affiliation	Isolation Source	ID (%)
Lewis Bac 1	MH329929	*Bacillus subtilis* strain NB-01 (HM214542)	Forest floor	100
Lewis Bac 4	MH329932	*Bacillus subtilis* strain K21 (JN587510)	Fermented soy food	100
Lewis Bac 18	MH329945	*Bacillus tequilencis* strain JO-17 (MF321840)	Saline desert soil	100
Lewis Bac 2	MH329930	*Cryobacterium* sp. (AB872307)	Cryoconite sediment	100
Lewis Bac 3	MH329931	*Bacillus safensis* strain CF4 (KY085985)	Corn rhizosphere	100
Lewis Bac 7	MH329935	*Bacillus pumilus* ZB13 (EF491624)	NA	96
Lewis Bac 16	MH329943	*Bacillus safensis* strain U41 (CP015610)	Lake Untersee	96
Lewis Bac 5	MH329933	*S. maltophilia* (LT906480)	Mouth	100
Lewis Bac 6	MH329934	*S. maltophilia* strain Nc 15MA-2(KP296212)	Mouth	100
Lewis Bac 13	MH329940	*S. maltophilia* strain 2681 (CP008838)	NA	100
Lewis Bac 21	MH329948	*S. maltophilia* strain 2681 (CP008838)	NA	100
Lewis Bac 8	MH329936	*Bacillus niabensis* strain G3-1-20 (KC494318)	Soil	100
Lewis Bac 10	MH329937	*Pn. taichungensis* strain 043(JN975184)	Plant root rhizosphere	100
Lewis Bac 14	MH329941	*Bacillus thuringiensis* strain 043 (KY323329)	Baltic sea	100
Lewis Bac 15	MH329942	*Bacillus horikoshii* strain 20a (CP020880)	Sediment	99
Lewis Bac 19	MH329946	*Lysinimonas* sp. (MG934620)	Glacier	100
Lewis Bac 20	MH329947	*Arthrobacter agilis* strain L77 (AY131225)	Psychrotolerant	100

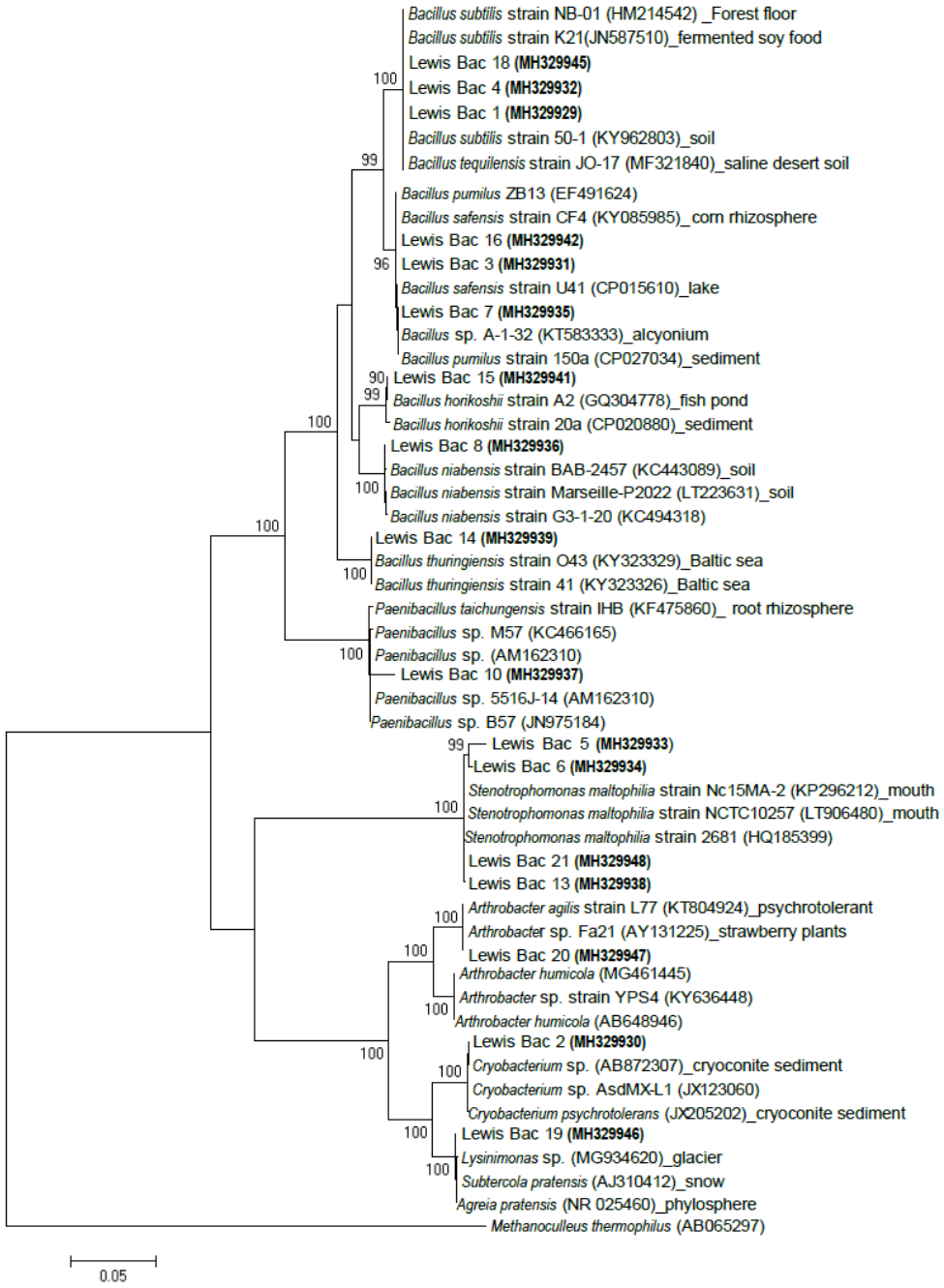


Figure 4.22: Phylogenetics of the snow isolates and some selected known bacterial species. *Methanoculleus thermophilus* (AB065297) was used to root the tree.

4.4. Composition and diversity of the microbial communities

ASV richness at 3% distance amounted to 993 in the two glacier sites (upper and lower). 451 ASVs were shared between the two sites while 680 ASVs were distributed in the upper site and 313 were distributed in the lower site (**Figure 4.23 a**). 3264 ASVs were distributed across the 23-year-old and 42-year-old rhizosphere samples. 774 ASVs were shared between the samples. The number of ASVs increased with the increase of age along the side-foreland chrono-sequence, 1626 ASVs occurred in the 23-year-old sample and 1638 ASVs occurred in the 42-year-old sample (**Figure 4.23 b**).

ASV richness between the glacier and the glacier foreland samples (cryoconite, rhizosphere and soil samples) mounted to 11,112 (**Figure 4.23 c**). Each of the sample types varied in the ASV distributions (rhizosphere = 6036, soil = 4240 and glacier =836 ASVs). Out of the 11,112 ASVs, 185 were shared across the three samples types, 2436 were shared between the soil and rhizosphere, 380 were shared between the soil and the cryoconite and 43 ASVs were shared between the rhizosphere and the cryoconite samples (**Figure 4.23 c**).

The ASV richness in the side-foreland soil samples (12-year-old, 23-year-old and 42-year-old soil samples) amounted to 2886 (**Figure 4.23 d**). Out of which, 34 were shared across the three samples, 430 were shared between the 23-year-old and 42-year-old samples, 35 were shared between the 12-year-old and 23-year-old samples while 6 ASVs were shared between the 12-year-old and 42-year-old soil samples (**Figure 4.23 d**).

Age also affected the ASV distribution along the glacier foreland rhizosphere and soil samples (**Figure 4.23 e, f**). In the rhizosphere, the ASV richness amounted to 4879. 271 ASVs were shared across the three sites (12-year-old, 23-year-old and 42-year-old rhizosphere samples), 348 ASVs were shared between the 12-year-old and 23-year-old samples, 454 ASVs were shared between the 23-year-old and 42-year-old samples while 100 ASVs were shared between the 12-year-old and 42-year-old rhizosphere samples

(**Figure 4.23 e**). The ASV richness in the glacier foreland soil chrono-sequence amounted to 4233. 351 ASVs were shared across the 12-year-old, 23-year-old and 42-year-old soil samples (**Figure 4.23 f**). 298 ASVs were shared between the 12-year-old and 23-year-old soil samples, 482 ASVs were shared between the 23-year-old and 42-year-old soil samples whereas 132 ASVs were shared between the 12-year-old and 42-year-old soil samples (**Figure 4.23f**).

Chao1 diversity index revealed a residual standard error of 79.07 on 71 degrees of freedom with a multiple R-squared of 0.842, adjusted R-squared of 0.8041, F-statistic of 22.25 on 17 and 71 DF with a p-value < $2.2e^{-16}$. The Chao1 analyses of variance (ANOVA) revealed a significant deference in bacterial community structure in all the sites (p-value <$2e^{-16}$). The taxa variation within the samples was revealed by the Chao1 Turkey comparison of means within the samples with p-value between $8.6e^{-6}$ to 1 (**Figure 4.24**).

Shannon diversity index revealed a residual standard error of 0.327 on 71 degrees of freedom with a multiple R-squared of 0.9061, adjusted R-squared of 0.8837, F-statistic of 40.32 on 17 and 71 DF and a p-value < $2.2e^{-16}$. The Shannon analyses of variance (ANOVA) revealed a significant deference in bacterial community structure in all the sites(p-value <$2e^{-16}$). The taxa variation within the samples was revealed by the Shannon Turkey comparison of means within the samples with p-value between 0.00 to 1 (**Figure 4.24**).

InvSimpson diversity index revealed a residual standard error of 38.88 on 71 degrees of freedom with multiple R-squared of 0.7247, adjusted R-squared of 0.6588, F-statistic of 11 on 17 and 71 DF and a p-value of $8.706e^{-14}$. The InvSimpson analyses of variance (ANOVA) revealed a significant deference in bacterial community structure in all the sites(p-value < $8.71e^{-14}$). The taxa variation within the samples was revealed by the InvSimpson Turkey comparison of means within the samples with p-value between $0.4e^{-6}$ to 1 (**Figure 4.24**).

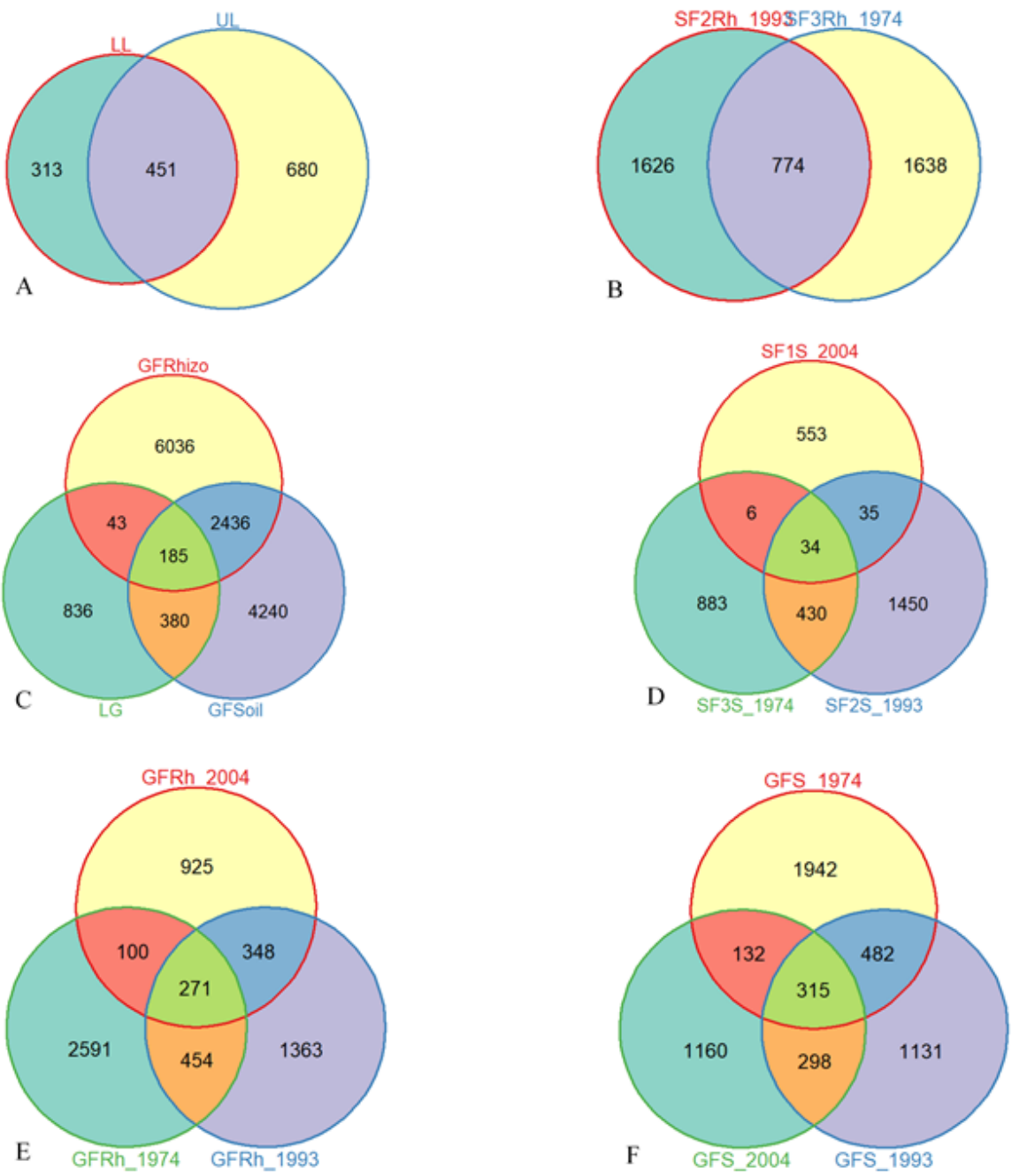


Figure 4.23: Shared ASVs between the two glacier sites, UL and LL (A), two side-foreland rhizosphere samples (B), glacier and two foreland rhizosphere and soil samples (C), side-foreland soil samples (D), glacier foreland rhizosphere (E) and glacier foreland soil samples (F).

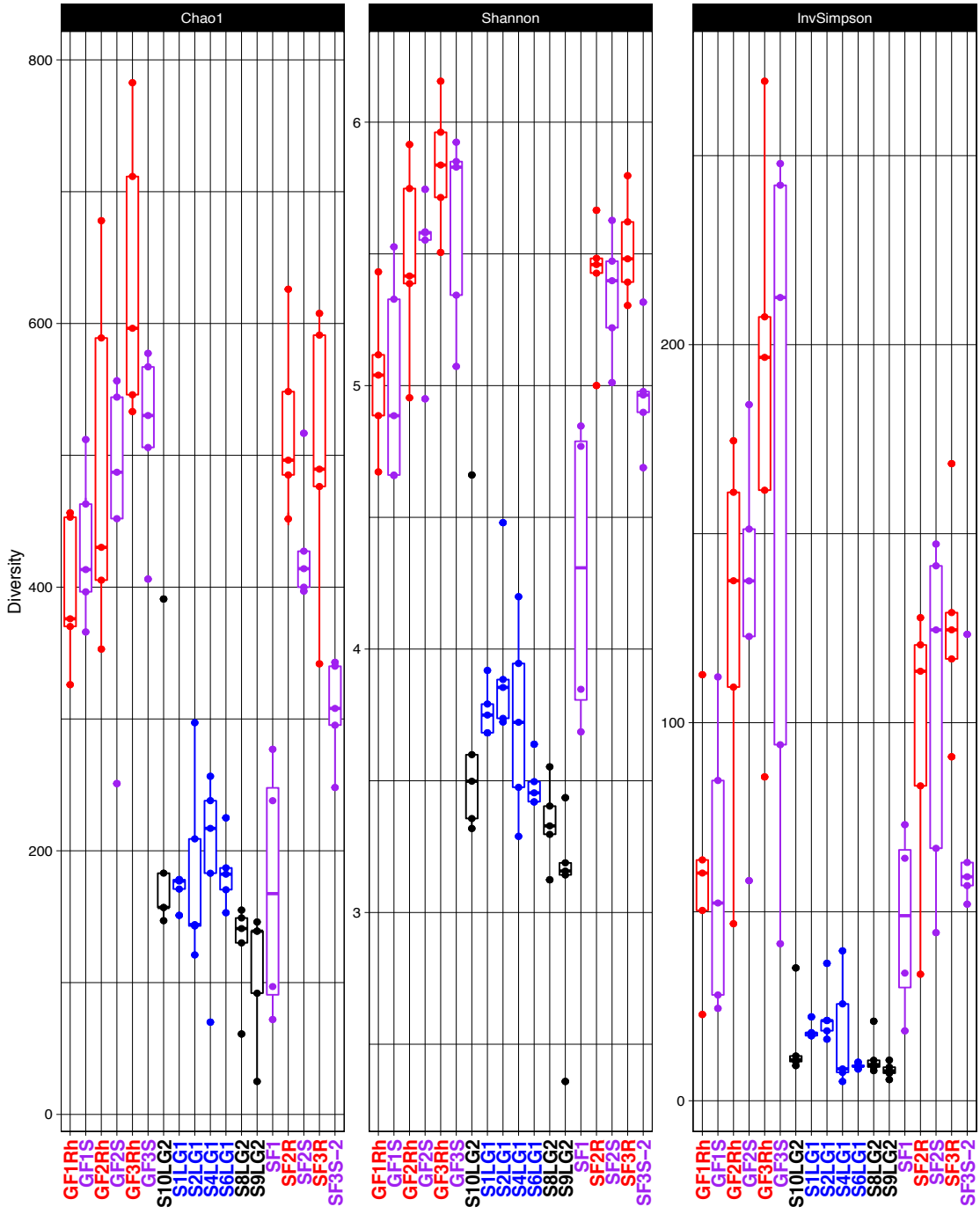


Figure 4.24: Alpha diversity indices within the glacier and foreland samples (red = rhizosphere, purple = soil samples, black = glacier lower site and blue = glacier upper site samples).

The rarefaction curves of 3% distance ASV analyses (**Figure 4.25**) indicated that, the rarefaction sequencing depth was far from being exhaustive in a number of samples even in the largest samples like GFRhizo_74. Only one sample out of the thirteen samples (SFSoil_04) reached a sequence depth plateau for exhaustive analyses.

The Principal Component Analyse (PCoA) and Non-metric multidimensional scaling (NMDS) analyses supported by ASV and taxonomic composition, divided the datasets into respective components (**Figure 4.26a** and **b)**. The samples were clustered according to sample sites and sample types. In the PCoA samples were grouped into four major clusters, but into two components as glacier (lower and upper sites) and foreland samples (foreland and foreland_Rhizo; **Figure 4.26a**). NMDS analyses grouped samples into three major clusters of soil (foreland), rhizosphere (foreland_rhizo) and glacier samples (upper and lower sites) as in figure **4.26b.** There was a major overlap of taxa in the upper Lewis and lower Lewis samples (**Figure 4.26b**). The overlap in the glacier samples was also supported with an insignificant difference in the paired t-test of sample means ($t = -0.41882$, $df = 28$, $p = 0.6785$).

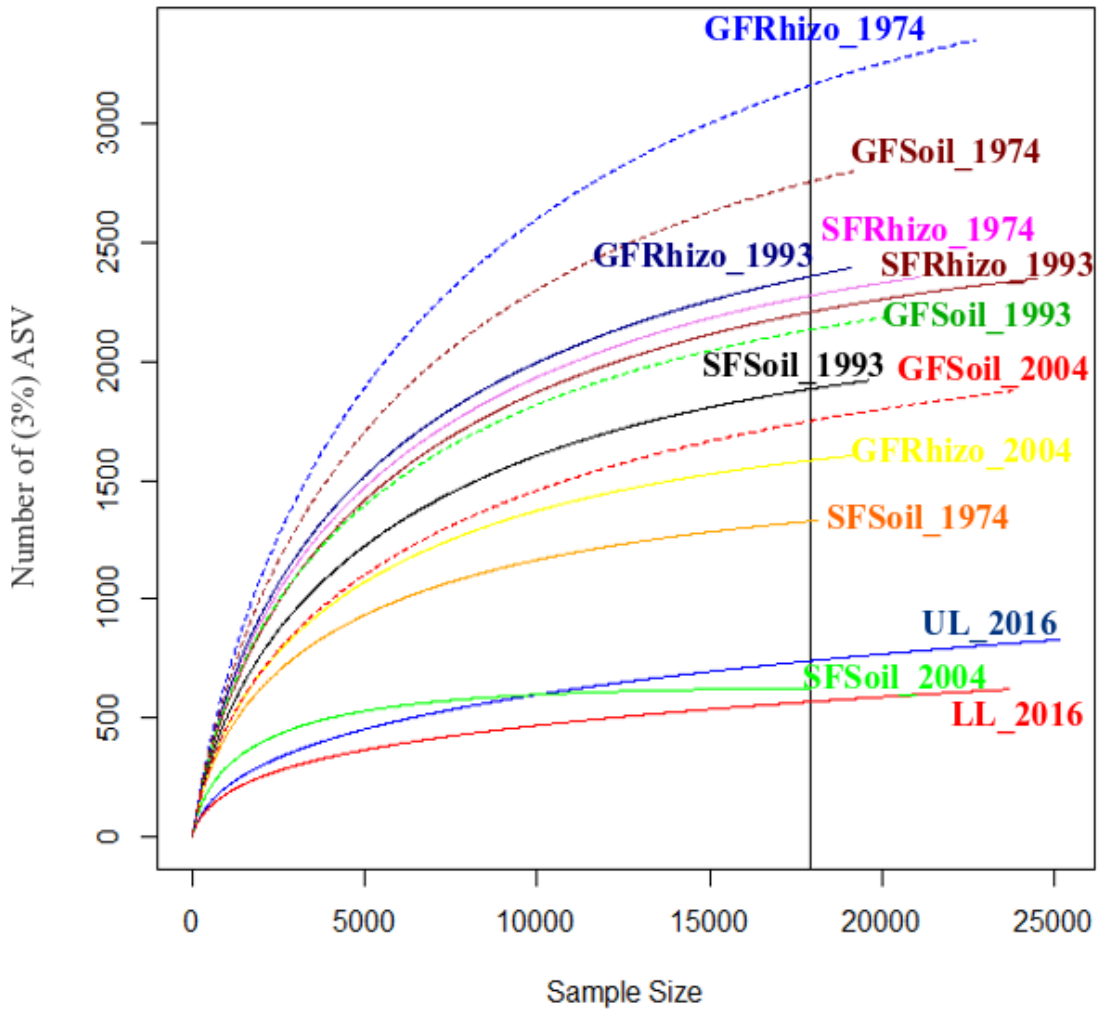


Figure 4.25: Rarefaction curves of ASVs in relation to sub-sampled sequence datasets size (number of reads). (UL = Upper Lewis, LL = Lower Lewis GFRh/Rhizo = Glacier foreland rhizosphere, GFSoil = soil samples, SF = side foreland, (_2004, _1993, _1974 = ages for the last glacier terminal).

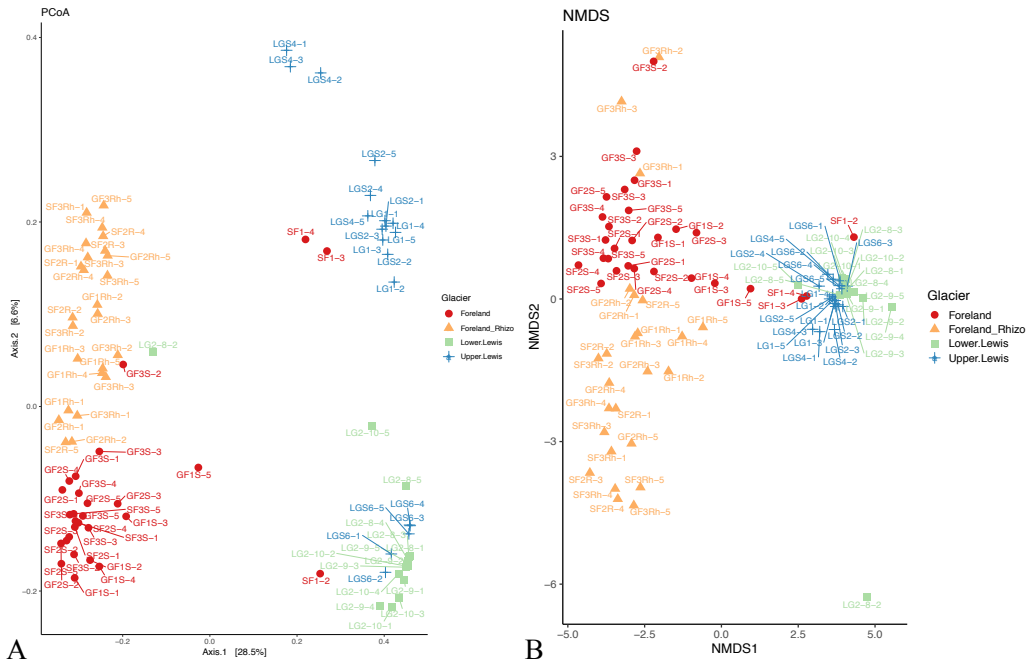


Figure 4.26: Bray-Curtis dissimilarities for PCoA and NMDS analyses between microbial compositions of 16S rDNA dataset grouped according to sampling sites. A: represent the PCoA, B: represent the NMDS, Foreland = soil samples.

The ordination of the mean squared contingency coefficient based on the Chi-squared distance and *cca* function in Correspondence analyses (CA) clustered the phyla according to sample type and sampling sites (**Figure 4.27a and b**). Most of the phyla averages were shared between the soil and rhizosphere samples that clustered together within a component to the right of the original axis. The phyla averages in the cryoconite samples were, however, clustered separately in the opposite component of the original axis. The rhizosphere samples showed possibility of high microbial community diversity as indicated by the blue sphere (**Figure 4.27a)**, which overlapped into the cryoconite samples**.**

The correspondence analyses of the mean squared contingency coefficient of the phyla in the sampling sites showed that, glacier sites (upper and lower) shared a number of phyla

and both the foreland sites also shared the phyla. The phenomenon was presented in two opposite components of the original axis (**Figure 4.27b).** Both the side-foreland (Sforeland) and the glacier foreland (Gforeland) showed a possible high diversity of microbial community structures, especially the side-foreland (**Figure 4.27b),** which overlaped to the opposite component occupied by the glacier samples.

In order to neutralise the long ecological gradient effect on the mean squared contingency coefficient of the Chi-squared distance across the glacier and foreland chrono-sequence in correspondence analyses, a Detrended correspondence analyses was also tested on the major phyla. The detrended analyses was based on the *decorana* function to support the down-weighted rare phyla. The variation in phyla diversity was significantly affected by the different sample types, cryoconite (p = 0.001), rhizosphere (p = 0.001) and soil (p = 0.02). The phyla were also clustered into groups supporting the detrended long ecological ecosystem **(Figure 4.28)** against the samples and the phyla.

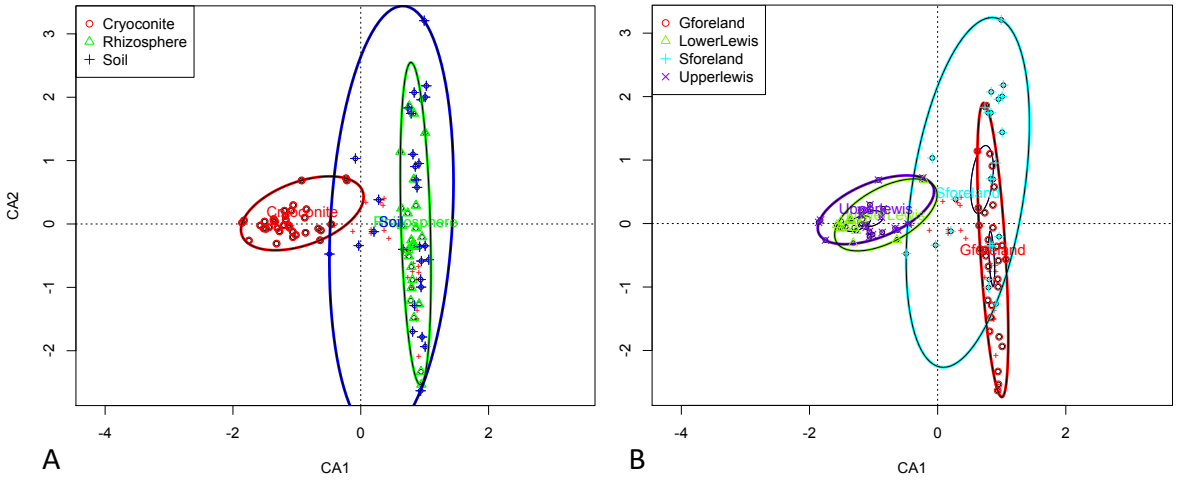


Figure 4.27: Unconstrained analyses of the relationship between the microbial community and the environmental constrains, based on the sample types (A) and sampling sites (B) based on the Euclidean method and *cca* function.

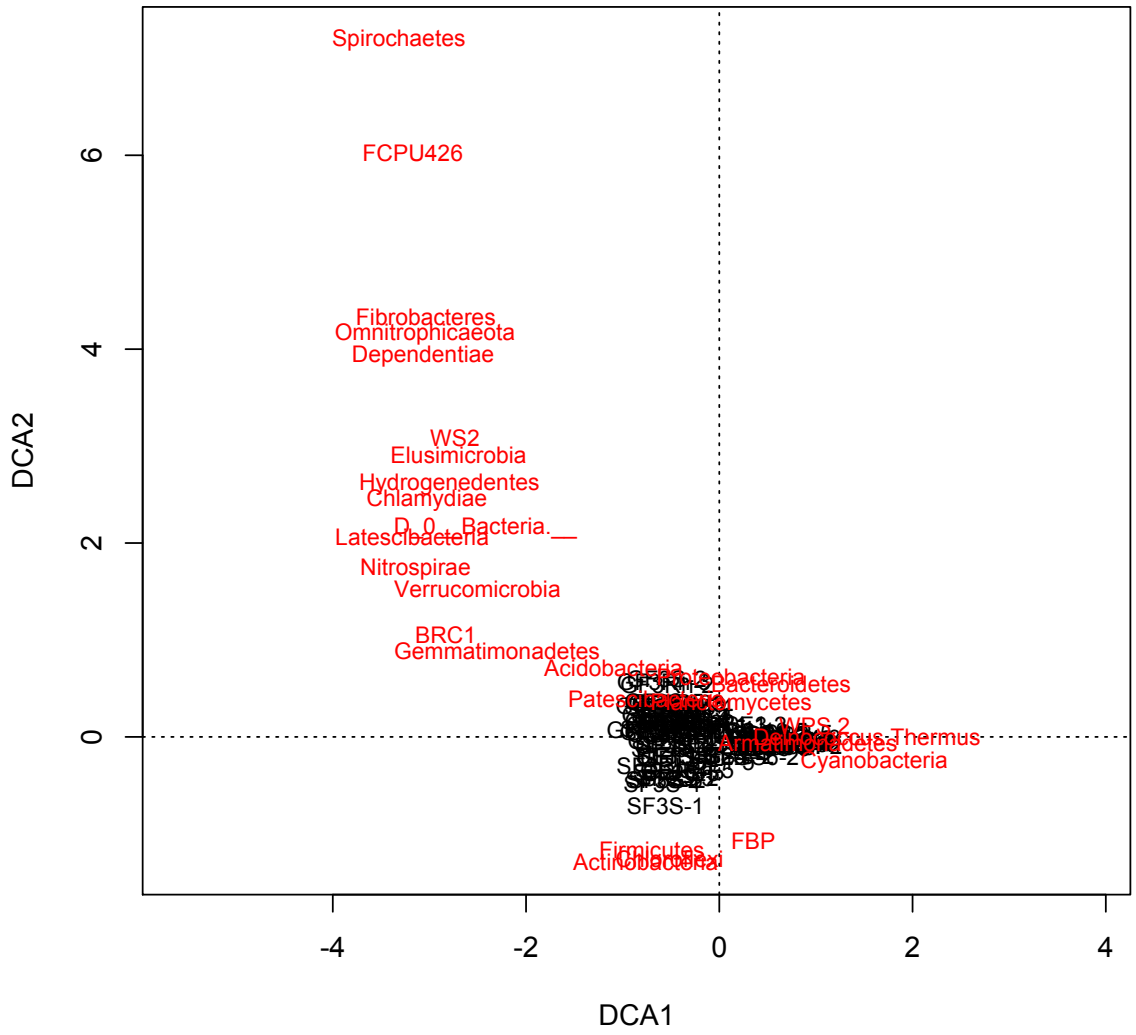


Figure 4.28: Detrended correspondence analysis (DCA) of the structure of major phyla in Lewis glacier and its foreland.

4.5. Indicator species in the glacier and Glacier foreland samples

The analyses of indicator organisms to determine the species that controlled the changes in the glacier environment, species preference and succession of the vegetational type significantly ($p < 0.002$) revealed 1 specie belonging to genus *Bradyrhizobium* to occur in the lower Lewis, 3 species to occur in the upper Lewis, 16 species to occur in the rhizosphere, 12 species occurred in the soil. Overall, 10 species occurred in the glacier foreland and 6 species occurred in the side foreland samples **(Figure 4.29)**. All these species affected the environmental changes and species preference across the glacier and its foreland.

Genus *Bradyrhizobium* was the indicator species on the lower Lewis with an indicator value of 0.6, species belonging to genus *Phormidesmis* was the indicator value on the upper Lewis with an indicator value of 0.9, species belonging to genus *Nakamurella* was the indicator species in the rhizosphere and it was closely followed by the species belonging to genus *Chthoniobacter* with indicator values of 0.9 and 0.8 respectively. Species belonging to genus *Acidiphilium* was the indicator for the soil with an indicator value of 0.8, species belonging to genus *gemmatimonas* was the indicator species in the glacier foreland with an indicator value of 0.7 while species belongning to genus *Blastocatella* was the indicator in the side foreland with an indicator value of 0.6. Genera were classified according to their phyla and different genera of were associated with specific sample types and sites (**Figure 4.29**).

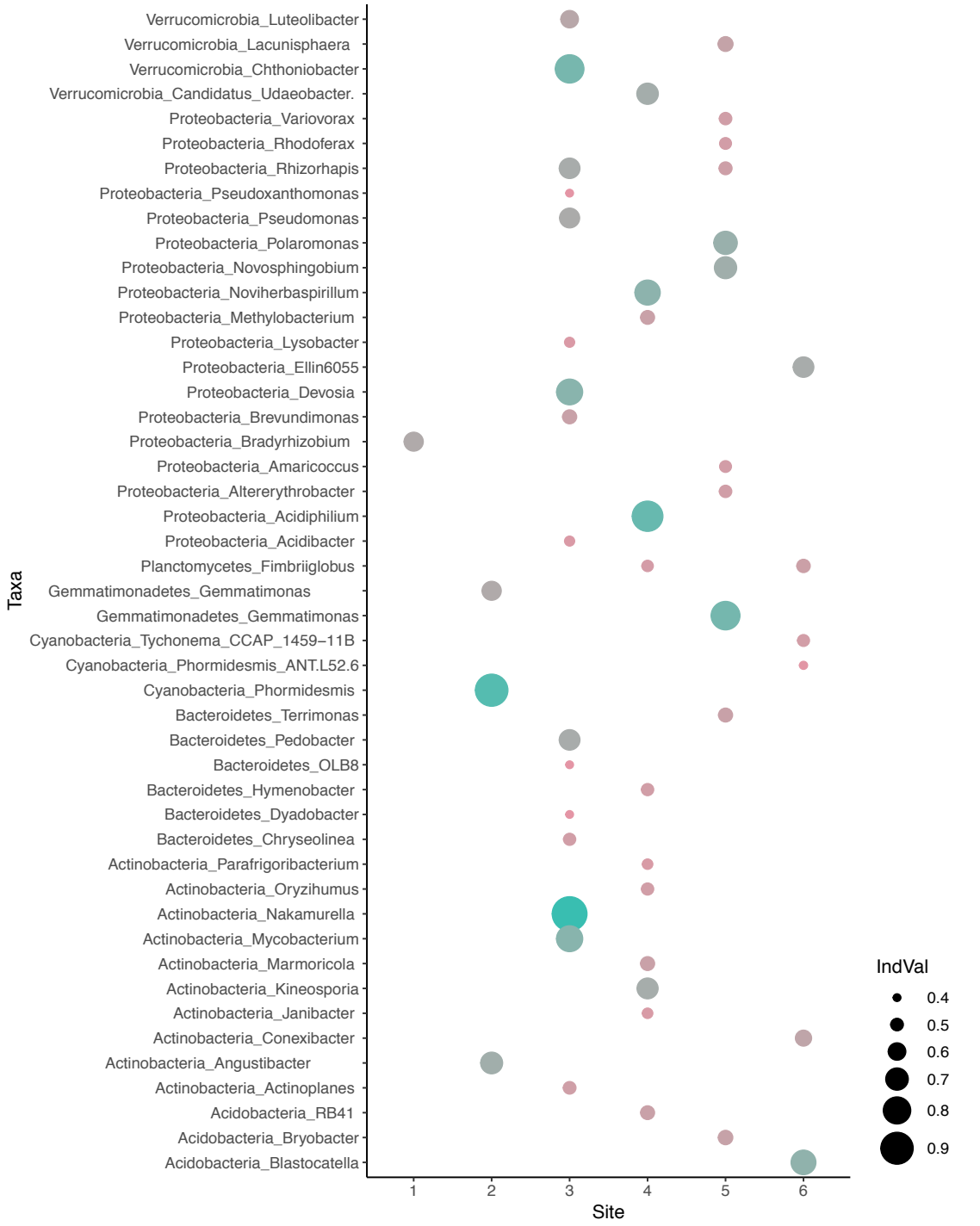


Figure 4.29: Indicator species on the Lewis glacier and its foreland at phylum level. Site (1 = Lower Lewis, 2 = Upper Lewis, 3 = Rhizosphere, 4 = Soil, 5 = glacier foreland, 6 = side foreland).

4.6 Relationship between the environmental factors and the microbial community structures on the Glacier surface

4.6.1 Correlation between the environmental constraints

There were six major factors that were considered to be of significance to bacterial distribution, abundance and community structure. These factors were cryoconite material, cryoconite organic matter, minerals, albedo, organic matter and altitude. The correlation coefficient varied among the parameters according to Spearman, Pearson and Kendall methods of correlation coefficient analyses. In all the three methods, cryoconite material and mineral were the most correlated factors (*r = 1*). They were followed by the cryoconite organic matter and organic matter (Spearman; *r = 0.9,* Pearson; *r = 0.9* and Kendall *r = 0.8*). Pearson method indicated a strong correlation between the chlorophyll.a and altitude (*r = 0.9*) as compared to Spearman (*r = 0.8*) and Kendall (*r = 0.7*). Other parameters were negatively correlated with albedo indicating poor correlation to either of the factors **(Figure 4.30)**. The correlation coefficient indicated that cryoconite material and mineral strongly affected the diversity and composition of bacterial community structures.

The correlation between the environmental variables and the bacterial community structure was almost linear with a histogram skewed to the right (**Figure 4.31, 32**). The significance difference in the relationship between the two matrices was strongly supported by the Mantel test (*r = 0.1763, p = 0.004*). The skewed histogram is probably due to the variation in taxa relative abundance across the glacier region. The differences between the prominent taxa and the less prominent taxa would also affect the strength of the correlation. The strength of the correlation value would also be affected by the distance differences which were performed based on the altitude and stake location. The dissimilarity based analyses of correlation between the environmental variable Euclidean distance and Bray-Curtis dissimilarity matrix from the bacterial community revealed altitude and Chlorophyll.a to strongly influence (*r = 0.318*) the community structure of bacteria in Lewis glacier. The significance of the two parameters was strongly supported

by the pairwise variation inflation analyses of predictors correlation (*r = 0.318*). altitude and Chlorophyll.a were closely followed by Altitude and cryoconite organic matter (*r* = 0.287). Chlorophyll.a was, however, the single best parameter that influence the taxa counts on the glacier surface (*r = 0.31*).

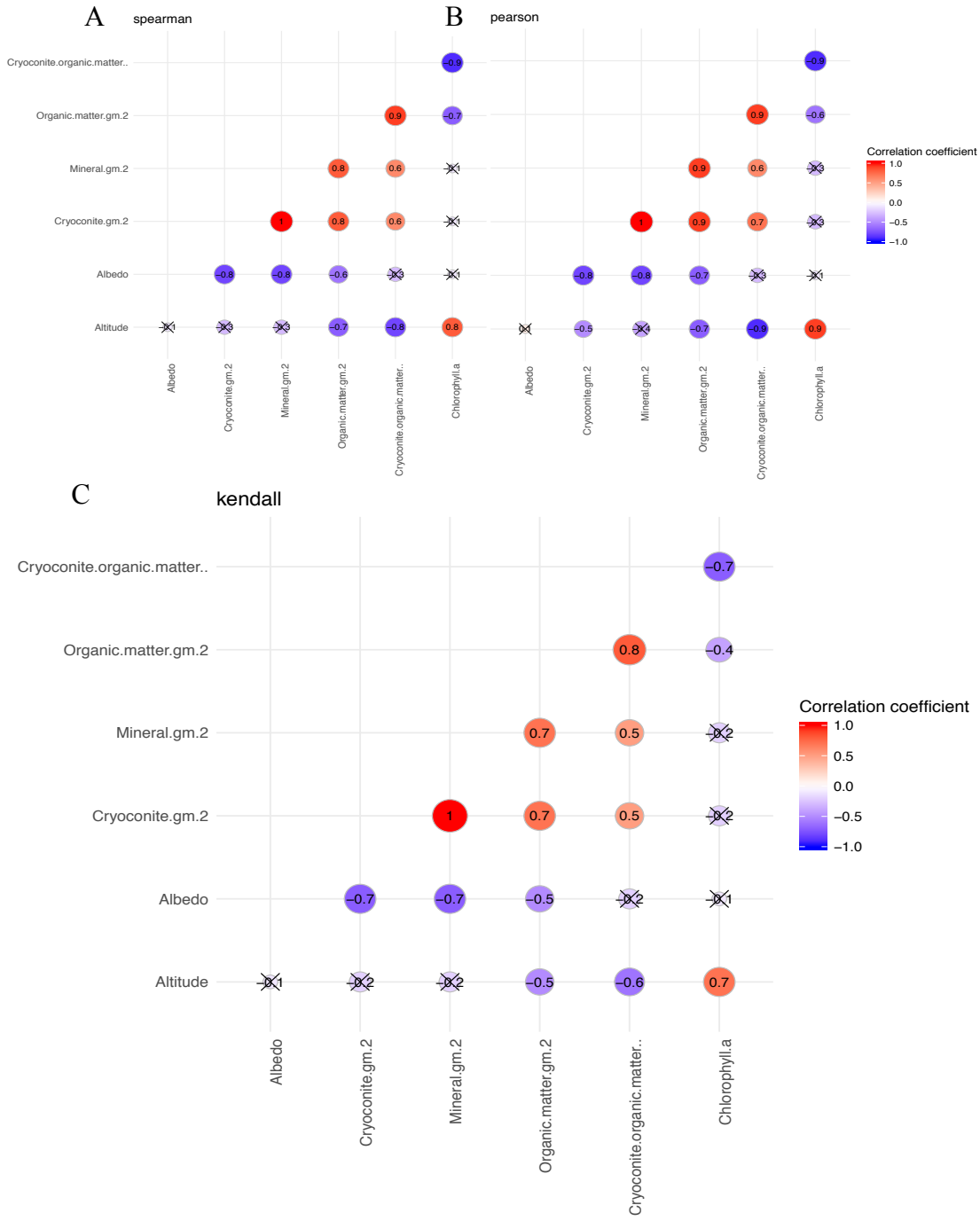


Figure 4.30: Correlation between the environmental constraints that had significant impact on the bacterial diversity on the glacier.

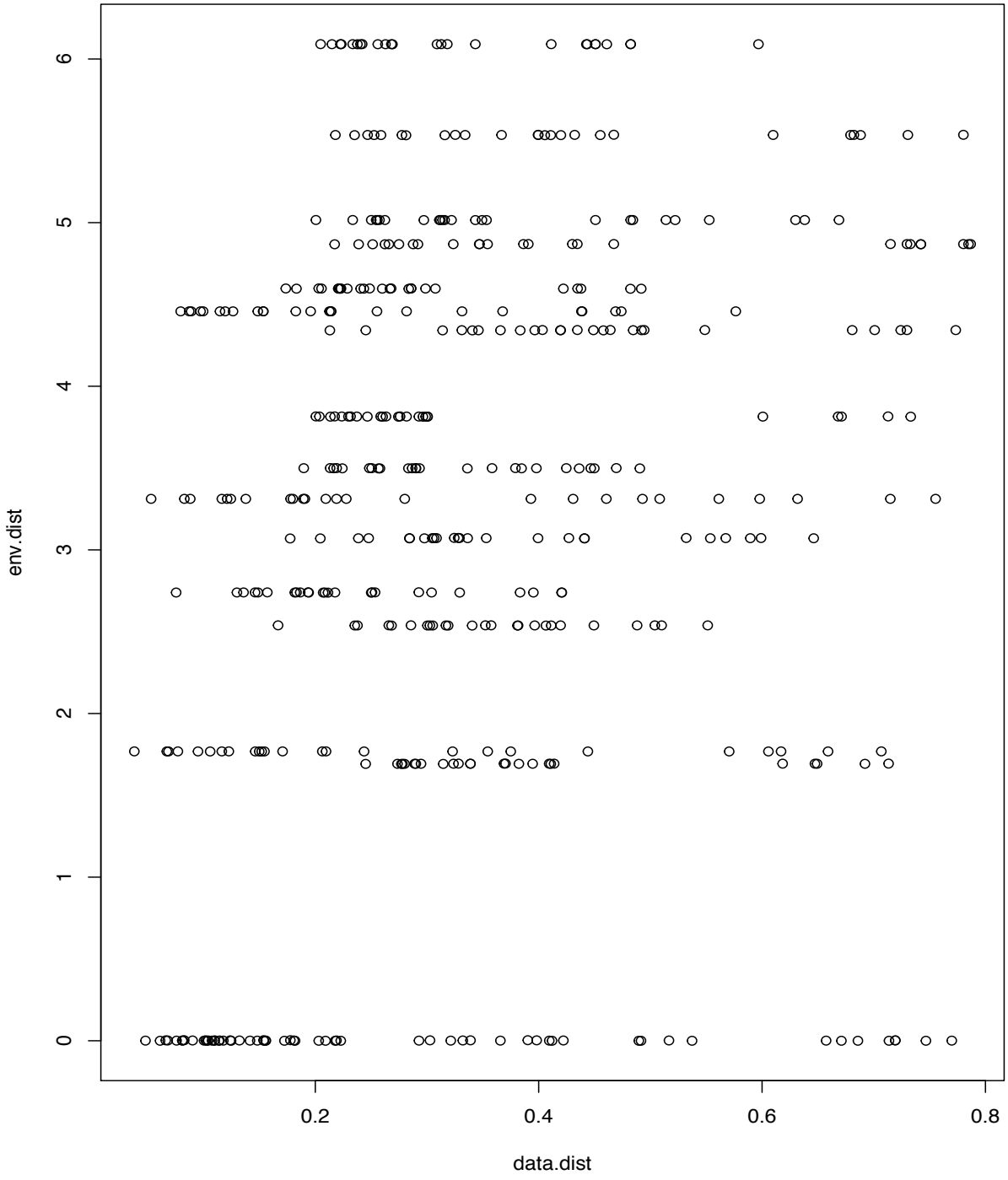


Figure 4. 31: Relationship between the bacterial community and the environmental matrices.

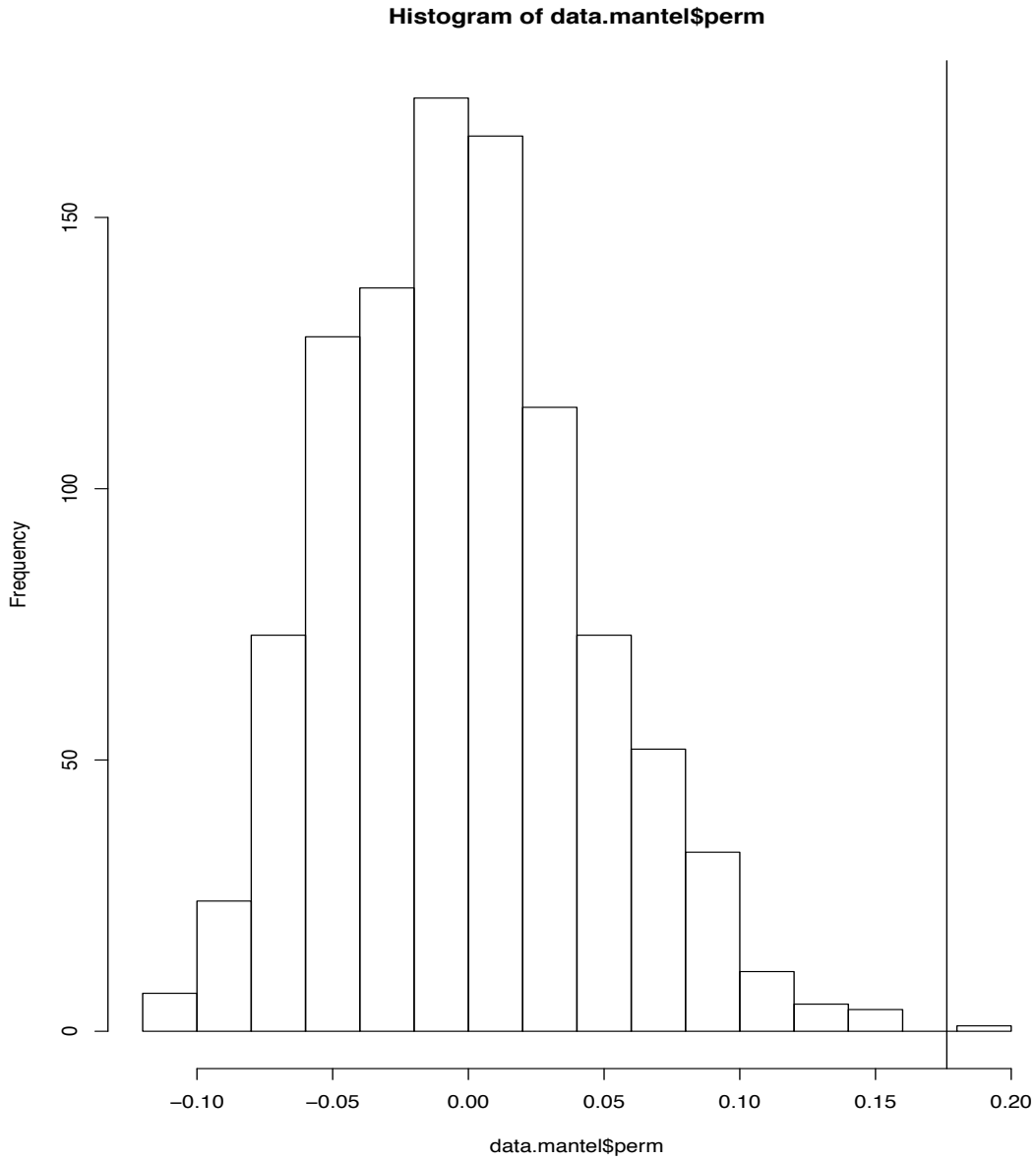


Figure 4.32: Histogram showing the significance of the Mantel test to bacterial community structure and environmental matrices.

The two factor order multivariate analyses of the six environmental factors showed that cryoconite organic matter was the best predictor of the bacterial diversity and community structure ($R^2 = 0.17, p < 0.0001$) against altitude and albedo. It was followed by altitude ($R^2 = 0.15, p < 0.0002$) against cryoconite organic matter and

albedo. Albedo did not show significance influence as a predictor against any of the variables ($R^2 = 0.04, p = 0.19$) against cryoconite organic matter and ($R^2 = 0.04, p = 0.202$) against altitude (**Table 4.2**).

Table 4.1: Results from two factor order PERMANOVA analyses on environmental factors, using Bray-Curtis distance on variance-stabilizing transformed counts.

Factor	R^2	p
Altitude + cryoconite organic matter	0. 15/0.04	<0.0002 ***/0.2191
cryoconite organic matter + Altitude	0.17/0.02	<0.0001 ***/0.6278
cryoconite organic matter + Albedo	0.17/0.03	<0.0002 ***/0.3194
Albedo + cryoconite organic matter	0.04/0.16	<0.1862 /0.0005 ***
Albedo + Altitude	0.04/0.14	<0.2016/0.0002 ***
Altitude + Albedo	0.15/0.04	<0.0002 ***/0.2875

The Constrained Analyses Proximities (CAP) between the environmental factors and the distribution of the prominent genera of *Cyanobacteria* indicated that, albedo, altitude, ch lorophyll.a, cryoconite material and organic matter affected the distribution of these gene ra (**Figure 4.33**). Figure **4.33** further showed that only genus *Tychonema* was never affec ted by the environmental factors. The gradient of the cryoconite organic matter indicates how this factor greatly affected the distribution of bacterial communities within the samp les. Cryoconite materials and albedo had least effect on the bacterial distributions across t he glacier surface (**Figure 4.33**).

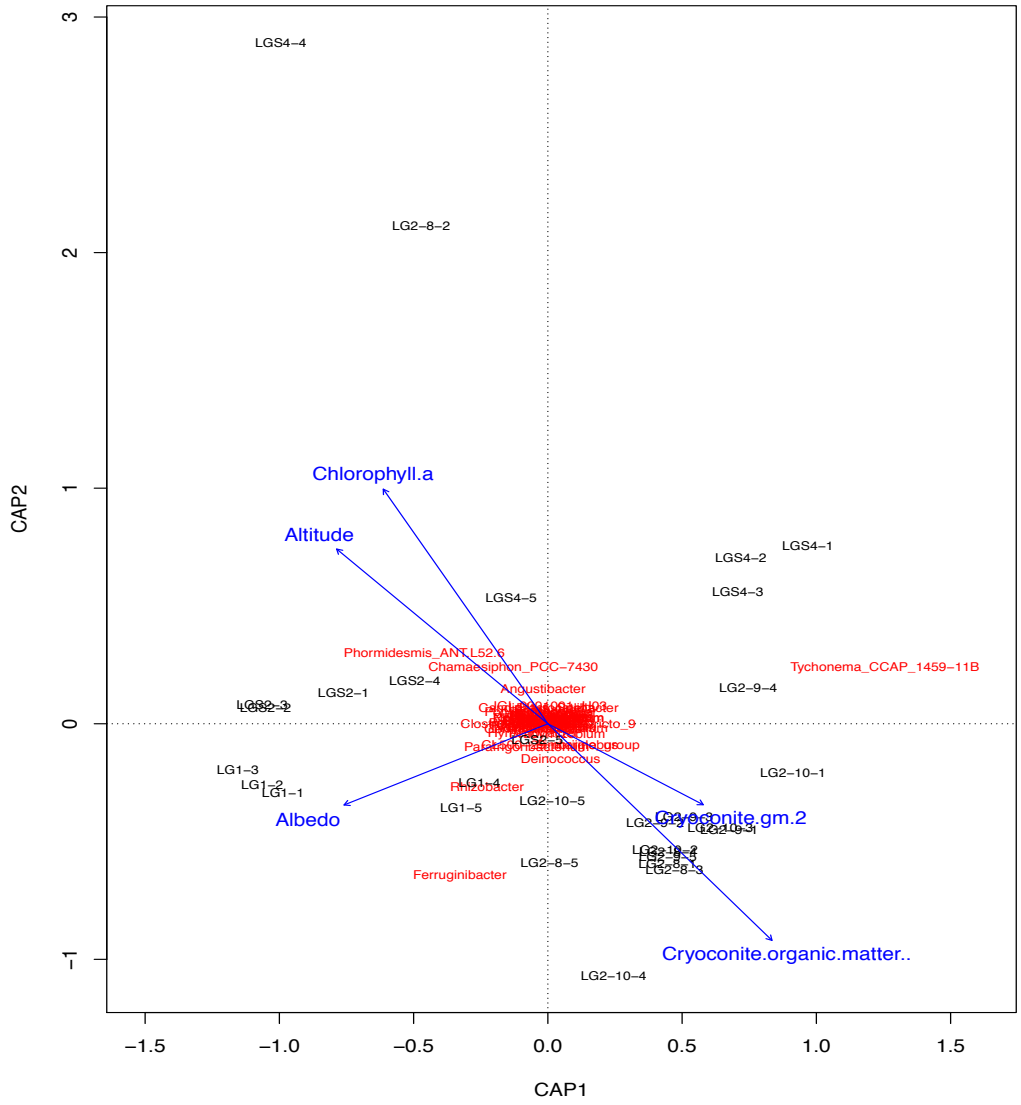


Figure 4.33: Constrained proximities between the environmental factors and bacterial community structures.

CHAPTER FIVE

DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

5.1 Discussion

5.1.1 Abundance of prokaryotic taxa in Lewis glacier and its Foreland

The relative abundance revealed 16 major phyla across the glacier and the foreland chrono-sequences. *Cyanobacteria* was the prominent and indicator phylum in the glacier, while *Proteobacteria* was the prominent phylum in the foreland chronosequence. However, *Gemmatimonadetes* was the indicator phylum in the foreland. The distribution of the phyla in the glacier sites is lower as compared to the high abundance of the community structure in the upper and lower sites of the glacier. In addition to *Cyanobacteria,* phyla *Protebacteria, Bacteriodetes* and *Actinobacteria* were also abundant with significant values of between 8% to 22% in the glacier sites.

The abundance of *Cyanobacteria* drastically reduced along the foreland chronosequence. The higher abundance in the foreland was observed in the primary foreland (12-year-old soil). This was different from the other dominant phyla, *Protebacteria, Bacteriodetes* and *Actinobacteria,* that dominated the foreland chronosequence with higher relative abundance with the increase of soil age.

In the primary foreland, *Cyanobacteria* is present as cyanotoxins, which are nutritious deposits from the glacier endogenous *Cyanobacteria* (Mur *et al.,* 1999; Makhalanyane *et al.,* 2015; Dulic *et al.,* 2016). Cyanotoxins are useful substances for the *Cyanobacteria* in their habitat for adaptability and host defence mechanism. In the foreland deposits, they are also useful to the exogenous colonizers from the aerial and rock influxes as nutritious substrates (Mur *et al.,* 1999). The availability of cyanotoxins in the primary foreland is important in supplementing the patchy resources (Tscherko *et al.,* 2003). These substrates control the abundance of other phyla in the glacier sites due to their active production from

the dominant *Cyanobacteria,* which are not active in the foreland ecosystem. In turn, there is exponential increase in the abundance of other phyla, that compete for resources.

Phyla *Protebacteria, Bacteriodetes* and *Actinobacteria* are very important in the preparation of the foreland for subsequent colonization by other organisms including the saprophytes, fungi and even the vascular plants (Schutte *et al.,* 2009). *Proteobacteria* are known for their mineralization capability (Yoshitake *et al.,* 2010). Elsewhere, the analyses of 16S rDNA has revealed the ability of *Proteobacteria, Firmicutes* and *Actinobacteria* to degrade organic compounds such as the propionate, acetate and formate that are available on the high altitude microbial mats (Hodson *et al.,* 2008). In this study, *Firmicutes* were, however, present in low abundance in the 12-year-old soil, and the major mineralization processes might have been taken over by the *Proteobacteria* and *Actinobacteria* that are prominent in the foreland chrono-sequence.

Phyla *Nitrospirae* and *Elusimicrobia* were more specific to the 42-year-old rhizosphere samples. They are the examples of the plant root-associated microorganisms (Davey *et al.,* 2015), that form symbiotic associations with the vascular plants. They play an important role in the directional replacement model (Walker and Del Moral, 2003) where primary colonizers of the foreland are replaced by the steady colonizers that form mycorrhizal association in the plant roots (Davey *et al.,* 2015). They also play a role in nitrogen fixation. This occurrence is also similar to the abundance of *Proteobacteria* and *Bacteriodetes* in the 23/42-year-old soil and rhizosphere samples while phylum *Actinobacteria* is more abundant in the 42-year-old soils and the rhizosphere of the side foreland. These three phyla were however, hierarchically clustered together in the heatmap clades.

This study revealed that, the dynamics of the microbial community structure in the foreland is species specific and not generalised to a specific phylum or kingdom of the prokaryote as explained by Nemergut et al. (2007). Moreover, the phenomenon may be

specific to certain regions based on the weather conditions and exposed activities from the immediate environment.

5.1.2 Abundance of the predominant phyla

5.1.2.1 Phylum *Cyanobacteria* in the Glacier.

Cyanobacteria is the prominent phylum in Lewis glacier with major genera such as *Phormidesmis, Microscoleuos, Leptolyngbya*, *Tychonema, Pseudanabaena, Chamaesiphon* and many more unknown genera. These genera, however, were distributed based on the sample sites. Some genera like *Tychonema* were sample type specific and was only clustered at the upper site. The distribution of the ASVs on the glacier surfaces might have been affected by the topographic orientation of the upper and the lower sites, which are relatively flat and sloppy respectively.

Most of these genera are filamentous mat-forming genera like the *Oscillatoria, Leptolyngbya, Phormidium* and *Nostoc*, that have been reported elsewhere to be significant in cryoconite hole stabilization and control (Stibal *et al.,* 2006). The unicellular genera that are affiliated to *Cyanobacteria* have received little attention regardless of their diversity in the glacier ecosystem (Boetius *et al.,* 2015). In this study some of the less reported genera like the *Oxyphotobacteria, Pseudanabaena* and *Chamaesiphon* were identified from the 16S rDNA analyses of the tropical samples. The presence of these genera in the highland cold regions of the tropical Africa like mount Kenya corroborates with the global analyses of the 16S rDNA. The study of 16S rDNA has revealed several glacial species, such as *Phormidium priestleyi*, *Leptolyngbya antarctica* and *Leptolyngbya frigida*, to be present in the arctic and Antarctic (Jungblut *et al.,* 2012). This phenomenon indicates a shared evolutionary history of colonization across the cold habitats.

Temporal changes also affect the distribution of the microbial communities on the glacial ecosystems. The significant difference in the distribution of the microbial communities between glacier site samples ($r = 0.256, p = 0.001$) was an indication of genera variation across the chrono-sequences. The significant difference was also evident with the impact

of altitude, age and sites to microbial community distributions across the glacier sites (Fig. 4.2.10). The changes are possibly attributed to the rapid loss of the channelized supraglacial materials into the subglacial avalanching processes (Hodson *et al.,* 2008) and influx of the exogenous factors.

The unknown *Cyanobacteria* genera were probably chloroplasts, which are identified as *Cyanobacteria* (Zarsky *et al.,* 2013). The two sites of Lewis glacier lack accumulation zone (Prinz *et al.,* 2012), which is an important region to hold the glacier content. This might be a factor to the low genera at the lower sloppy site of the glacier and high number of genera at the relatively flat upper site of the glacier.

Genus *Cyanobacteria* associated with the cryoconite produce cohesive extracellular polymeric substances compounds, which cause the aggregation of cryoconite materials (Takeuchi *et al.,* 2001, Hodson *et al.,* 2010, Langford *et al.,* 2010). Formation of the aggregates by the *Cyanobacteria* species is important in the prolonged residence time and creation of a microenvironment for other microorganisms (Bøggild *et al.,* 2010, Irvine-Fynn *et al.,* 2011). The prolonged residence time and large microbial aggregates also improve the metabolic processes within the glacier samples (Hodson *et al.,* 2010, Langford *et al.,* 2010).

The schematic orientation of the glacier surfaces is, however, known to facilitate the process through avalanching and crevassing of the glacier content (Hodson *et al.,* 2008). Crevassing enhances glacier basal sliding and till deformation (Paterson, 1994), which couples the supraglacial and subglacial ecosystems through hydrological transfer within the glacier ecosystem (Hodson et al., 2008). The slope of the glacier surface is suggested elsewhere (Zarsky *et al.,* 2013) to affect the development of the cryoconite aggregates due to erosive action of the meltwater. The phenomena affect the bacterial residence time and colonization of the specific sample sites within the glacial ecosystem.

5.1.2.2 Phylum *Cyanobacteria* in Glacier foreland.

The primary deglaciated foreland (0-year-old) of Lewis glacier has high diversity of five genera of phylum *Cyanobacteria*. Some of the primary foreland colonizers were clustered to have high community structure at the lower Lewis glacier. Most of the genera in the primary foreland are the possible nitrogen and carbon fixing species, that prepare the foreland chronosequence for the subsequent diverse microbial colonizers together with other symbiotic organisms. The success of *Cyanobacteria* species activity at the primary foreland is coupled with the association to other symbiont organisms like moss in the high altitude ecosystems (Gavazov *et al.,* 2010). Generally, the analyses of similarity revealed that, the diversity of *Cyanobacteria* is affected by the sample type, sample site, altitude and the age of the last glacier terminal ($r = 0.469, p = 0.001, r = 0.256, p = 0.003, r = 0.237, p = 0.019$ and $r = 0.569, p = 0.019;$ respectively).

Chamaesiphon is the only genus that is not affected by the directional and non-directional succession patterns as it is distributed across the soil and rhizosphere samples along the foreland chrono-sequence. The dominance of genus *Crinalium* community structure at the 12-year-old soil and genus *Tychonema* at the 42-year-old rhizosphere samples with a steady community structure is an indication of specialization of the foreland chronosequence by specific microbial communities. Some of the species of the genus *Microcoleuos* has been reported to colonize desert soils (Garcia-Pichel *et al.,* 2001), where they improve soil stability by the production of exopolysaccharides that aggregate soil particles at the high elevations of Arctic and Antarctic (de Caire *et al.,* 1997; Hodkinson *et al.,* 2003; Wynn- Williams, 2000). The microbial succession pattern at the glacier foreland suggests a gradual change in the microbial community structures and diversity. The primary foreland is colonised by some glacier specialists, which are deposited by the glacier tillage at the terminal melt points as debris from the endogenous glacier ecosystem. The colonization of the young foreland soil by the glacier specialists has been explained elsewhere (Nemergut *et al.,* 2006; Edwards *et al.,* 2014) as a useful process by the nitrogen fixing species for the subsequent

microbial community structures, bryophyte, lichen and vascular plant colonization. The colonization of the primary foreland by the endogenous glacier specialist is also supported by results from the Damma glacier catchments habitats, that revealed endogenous glacier debris as the major source of primary foreland bacterial colonizers (Priscu *et al.,* 1998; Rime *et al.,* 2016). Moreover, Hodkinson et al. (2003) suggested that *Cyanobacteria* are important to Carbon and Nitrogen fluxes in the newly deglaciated foreland soils. Genus *Tychonema* is therefore classified as the generalised colonizers, that inhabits the glacier and foreland rhizosphere ecosystems. Alternatively, this genus might be a nitrogen fixing that form symbiotic associations to other organism.

5.1.2.3 Proteobacteria in the Glacier and its foreland

The diversity of the major genera of the phylum *Proteobacteria* across the glacier and it foreland were affected by the sample type and age of the last glacier terminal in the foreland. The lower glacier site had high diversity of the ASVs possibly due to the exposure to the primary foreland and the moraine influxes from the upwind and the Indian ocean weather systems. The lower glacier site may also get a lot of microbial deposits from the surrounding falling rocks and stones. Generally, the lower site is rather steep compared to the upper site and due to this it is expected to have low diversity of microbial genera due to rapid flush of the glacier endogenous systems (Walder and Fowler, 1994; Paterson, 1994; Hodson *et al.,* 2008; Hodgkins *et al.,* 2013).

The low diversity of genera in the upper site of the glacier is important in the understanding of the dynamics of the glacier specialist genera and glacier generalised colonisers. The prominent genera in the upper glacier, *Rhizobacter* and *Rhodovastum* might be glacier specialists since they dominate the upper stable site and poorly distributed in the foreland chronosequence. Genus *Rhizobacter* inhabited the glacier sites, the 12-year-old foreland soil and the 23-year-old foreland soil while genus *Rhodovastum* inhabited only the 12-year-old foreland soil at lower relative abundance of 6.7 %

compared to the 30 % in the glacier. The difference in the colony preference indicate a probable environmental factor that influences the community structure of such genera.

Foreland of Lewis glacier demonstrated high diversity of the *Proteobacteia* genera. The diversity was, however, affected by the age and the sample type of either soil or rhizosphere. Genera diversity decreased with age from the 12-year-old foreland soil and rhizosphere to 42-year-old foreland soil and rhizosphere. Their occurrence indicated the stability of the community structure to increase with the increase of age down the foreland. At the 12-year-old foreland the genera diversity is high due to the heterogeneity and geochemical activities of glacier endogenous species and exogenous species from the air, moraine and rocks (Nemergut *et al.,* 2006; Yoshitake *et al.,* 2010; Edwards *et al.,* 2014; Rime *et al.,* 2016). The biogeochemical activities by the microorganisms in the primary foreland are important for the succession of subsequent soil and symbiotic colonizers (Rime *et al.,* 2016), including the vascular plants. The biogeochemical processes at the primary foreland also boost the patchy resource distribution (Tscherko *et al.,* 2003) which is replacement with time as the foreland soil incorporates new associations with species turnover (Welden and Slauson, 1986; Neff *et al.,* 2002).

The species turnover is defined by the directional replacement model of the primary colonizers through a competitive interaction between microbial community structures (Walker and Del Moral, 2003), while the directional non- replacement model leads to a progressive change and persistence taxa occurrence (Vater and Matthews, 2013) in soil and the plant rhizosphere. The directional non-replacement model facilitates the stability of the mycorrhizal organisms, especially the plant root-associated microorganisms (Davey *et al.,* 2015). This finding is in agreement with the current finding where the microbial diversity reduces down the foreland age, but the species community structures are stabilised within a few genera in the 42-year-old foreland soil and rhizosphere.

5.1.2.4 Phylum Bacteroidetes in the Glacier and its foreland

The genera affiliated to phylum *Bacteroidetes* occurred in both the glacier and the foreland chrono-sequence. The genera were diverse in the foreland than the glacier samples. However, their diversity in the foreland was dominant in the young soil samples than in the rhizosphere and in the old soils. Low abundance of phylum *Bacteroidetes* in the glacier ecosystem is not only limited to the tropical Lewis glacier of Mount Kenya, since the low abundance has been recorded in other polar regions of the North Greenland Eemian Ice (Miteva *et al.,* 2009). Consequently, the abundance of *Bacteroidetes* in the sea ice and glacier ecosystem has been reported to occur with relatively higher percentage as compared to other known glacier specialists like phylum *Cyanobacteria* (Boetius *et al.,* 2015; Anesio *et al.,* 2017). Some of the major genera to this phylum, like *Flavobacterium* that dominated the foreland chrono-sequence of Lewis glacier have been reported to dominate alpine glaciers, where they express high cold-active hydrolases enzymes (Singh *et al.,* 2014).

The high diversity of the major genera affiliated to phylum *Bacteroidetes* in the young foreland soil of Lewis glacier significantly indicates the known role of such genera in other foreland ecosystems. The primary foreland is usually affected by the patchy resource distribution (Tscherko *et al.,* 2003), which is replaced with time as the foreland soil incorporates new associations with rapid species throughput (Neff *et al.,* 2002) to incorporate diverse microbial community structures, including affiliates to phylum *Bacteriodetes.*

The genera affiliated to *Bacteroidetes* significantly contribute to the directional replacement of the primary colonizers of the deglaciated foreland. Their occurrence in the young soil is further affected by horizontal replacement by the species that respond to the non-directional replacement model (Svoboda and Henry, 1987; Matthews, 1992; Walker and Del Moral, 2003; Vater and Matthews, 2013), since their diversity gradually changes with increasing age down the glacier foreland chrono-sequence. The microbial community

dynamics in the foreland is affected by the overtime shifts from the active biogeochemical processes in the primary foreland to steady recolonization of the bare land with vascular plants and mycorrhizal associations, which eventually create a stable ecosystem with high vegetation cover (Bardgett *et al.,* 2007; Schutte *et al.,* 2009; Davey *et al.,* 2015).

5.1.2.5 Phylum Actinobacteria in the Glacier and its foreland

The diversity of the major genera of the phylo-type *Actinobacteria* were influenced by the sample types and the locations across the glacier and the foreland ecosystems. The diversity was also influenced by the specificity and specialization by the major genera. The major genera could be classified as specialized glacier colonisers or as the generalised colonisers. The colonization strategy, however, influenced their ways to the glacier or to the foreland through the aerial or exogenous depositions.

The relative abundance of the major genera of the phylum *Actinobacteria* indicated low diversity of the genera in the glacier sites, but high community structures of the few genera in the same samples. The prominence of genera *Frigoribacterium* and *Oryzihumus* indicate a possible specialization of specificity to psychrophilic ecosystems while the presence of the marine groups of the phylotype *Actinibacteria* indicate the possibility of community influxes from the ocean weather or aerial microbial community deposition.

The primary foreland soil and rhizosphere were dominated by diverse genera: *Gaiella, Kineosporia, Pseudarthrobacter, Rhodococcus, Janibacter, Oryzihumus, Knoellia, Nakamurella.* Some of these genera, *Oryzihumus* and *Nakamurella* were observed to colonise the glacier sites. The abundance of the genus *Oryzihumus* steadily decreased from the glacier sites to the foreland while the abundance of genus *Nakamurella* was relatively dynamic due to the directional replacement of the glacier endogenous and primary colonizers through a competitive interaction between microbial community structures (Walker and Del Moral, 2003). The dynamics is also due to a progressive change and persistence taxa occurrence (Vater and Matthews, 2013) in the plant rhizosphere. The species turnover fluctuated at the 23-year-old rhizosphere and the community structure

stabilized at the 42-year-old rhizosphere. A similar trend is observed with the genus *Kineosporia,* that demonstrate dynamic abundance in the rhizosphere samples just as the genus *Nakamurella.* The phenomenon was, however, different in the soil samples, where the abundance of genus *Kineosporia* steadily increased with the increase of the soil age.

Generally, stability for microbial communities and structures is due to the adaptability and homogenization of resources within an ecosystem with and other mycorrhizal organisms. This would explain the common dynamics in rhizosphere samples where microbial heterogeneity and resource patchiness in the primary foreland (12-year-old) is succeeded by resourceful and symbiotic associations with the vascular plants and other symbionts in the subsequent soil ages. In the mid-age, 23-year-old, there is a possible competition between the symbionts leading to a stagnation and reduction in the number of microbial communities. The microbial communities, exponentially, increase at the 42-year-old samples where the resources are numerous and the mycorrhizal association is specified.

The foreland succession strategy by the microbial communities is somehow different in the soil samples as observed in a number of soil specialist genera. Genus *Knoellia* only colonized the soil samples, where its abundance progressively increased by the increase in the soil age. The same trend was observed in other genera like *Pseudarthrobacter*.

Competence in the primary foreland is a key factor in the success to succession mechanisms in the foreland ecosystems where the less competent microbial community structures are replaced by the most competent organisms. Genus *Janibacter* only colonized the primary foreland soil, where its abundant was relatively high. Also genus *Janibacter* was hierarchically clustered independently from other genera. Most probably, *Janibacter* is less competent to achieve the succession dynamics in the foreland ecosystem. Alternatively, it might be an exogenous coloniser that was outcompeted by the glacier endogenous communities that dominated the primary foreland.

Other genera, like *Rhodococcus* would be described as opportunistic colonizers, that are only observed in a specific ecosystem at a specific time. Genus *Rhodococcus* was only observed in the lower glacier site with lower relative abundance.

5.1.3 Prokaryotes in the supraglacial snow on Lewis Glacier

Snowpack of Lewis glacier is a rich psychrophilic ecosystem that is conducive for the microbial inoculi. The snow cover of Lewis is, however, experiencing the downwind bio-aerosol input from the surrounding tall rocky points, moraine, terrestrial, human activity and savannah forests from Laikipia county. In this study, the isolates were partially identified to be closely related to the human gut pathogens, terrestrial and psychrophilic microbial ecology. The molecular analyses of 16S rDNA gene sequences revealed three phyla, *Firmucutes*, *Proteobacteria,* and *Actinobacteria* as the colonies on the Lewis glacier snowpack.

The presence of the three phyla *Firmicutes, Proteobacteria,* and *Actinobacteria* on the snowpack would be an indication of active degradation of organic compounds from the glacier surface as compared to their known features on glaciers elsewhere (Hodson *et al.,* 2008). The occurrence of these phyla is also characterized with the balance of nutrient cycling between the atmosphere and glacier surface. Elsewhere Jones (1999), Hodson et al. (2005, 2008) have shown that nitrogen cycling is important to the melting polar glaciers. This would be much faster in the tropical glaciers, which are actively exposed to photosynthetic and heterotrophic processes in addition to other biogeochemical processes due to available litter and bio-aerosols influxes. However, the biogeochemical role of snowpack biota in nutrient cycling is not well understood.

The succession process of the rapidly disappearing Lewis glacier also contributes to the litter input on the snowpack. The genus *Bacillus* was the most dominant group on the glacier snowpack with a percentage of 53%. This can be explained by the strategic location of the glacier site to the windward and leeward sides where it experiences the upwind predominant weather systems from Indian ocean and downwind influxes from the

savannah zones of Nanyuki respectively. The exposure of the glacier surface to continuous leeward dust flow makes it a recipient settling zone for the contaminating terrestrial bio-aerosols. Most of the analyzed isolates (12; 53%) are soil and plant root rhizosphere symbionts while 4 (23.5%) are closely related to human gut pathogens. Probably, the symbionts would be easily blown by the downwind from the savannah zones and oceanic upwind from the glacier foreland, which is colonized by the vascular plants (Schutte *et al.,* 2009; Davey *et al.,* 2015) onto the snowpack that acts as a settling point for litters and bio-aerosols. Moreover, soil inoculants would arise from the glacier terminal moraine that is re-colonized by the snow generalist phyla. The phylum *Proteobacteria* is a snow generalized colony, which is known to play a significant role in the rock mineralization at the glacier foreland (Yoshitake *et al.,* 2010) that extends to the moraine ecosystem.

Isolates Lewis Bac 5, 6, 13 and 21 were identical to known members of the human pathogen (genus *Stenotrophomonas*) indicating that human activities on the glacier and also the surrounding influxes are possible sources. *Stenotrophomonas maltophilia* is ubiquitous in aqueous environments, soil, and plants (Berg *et al.,* 1996, 1999), which are great sources to possible contaminant on the Lewis glacier snowpack. They are also known to be useful in wide range of biotechnology applications (Bhattacharya *et al.,* 2007; Ryan *et al.,* 2008). However, their occurrence in cryophilic environment is a rear finding that is reported for the first time from the tropical glacier.

In this study only three snow specialist species (17.6%; isolates Lewis Bac 2, 19 and 20) were partially identified out of the 17 isolates. These three isolates were closely related to known groups of glacier specialist species of the genera *Cryobacterium, Subtercola/Agreia* and *Arthrobacter* that have only been published from the mid-latitude to polar regions (Hodson *et al.,* 2008). The genus *Cryobacterium* is psychrophilic (Suzuki *et al.,* 1997; Zhang *et al.,* 2007). *Cryobacterium psychrotolerans* is aerobic bacterium isolated from the China No. 1 glacier. They grow well between 4 – 27^{o}C with an optimum growth at 20 – 22^{o}C (Zhang *et al.,* 2007). These conditions can as well be inhabited in the tropical Lewis glacier, which seems to have a constant temperature conditions.

Arthrobacter agilis is a psychrotrophic bacterium, which occurs in lake water and Antarctic sea ice (Bowman *et al.,* 1997; Deming, 2002). It produced dimethylhexadecylamine, plant growth promoting enzymes and cold active hydrolytic enzymes (Nadeem *et al.,* 2013). These elements would be useful for the species competence, degradation of bioactive elements and cold shock stability within the tropical glacier ecosystem. Due to the seasonal snow fall on Lewis glacier and rapid melt of the englacial and subglacial zones of Lewis glacier (S. Hastenrath, 1983b), 2006; (Prinz *et al.,* 2011b), these psychrophilic prokaryotes are endangered and can be easily lost through the glacial surface run off and melt water tillage to the glacier bed and melting points.

Generally, the stability of microbial interactions on the snowpack is complexed with the climate change, duration of snow melt and rainfall (Hodson *et al.,* 2008). However, the snowpack layer of Lewis glacier is relatively thin and have shorter melting durations. Most likely due to the location of the glacier closer to the equator with a constant temperature. It means that the phototrophs are not affected by the snow flurries and their interactions with the snow heterotrophs have impact to the glacial interface microbial community structures, which may facilitate the rate of biogeochemical processes leading to rapid glacier melt. Lewis glacier snowpack is not colonized by the phototrophic prokaryotes as usual to other glaciers worldwide (Stanier and Bazine, 1977; Harding *et al.,* 2011). This might be due to the seasonal snow fall characteristic of the Lewis snow cover, which is coupled with the influx of the bio-aerosols.

The phototrophic prokaryotes might be colonizing the submerged cryoconite holes on the Lewis glacier surface, but not on the snowpack. The findings in this study suggest that the snowpack of the tropical Lewis glacier is colonized by diverse prokaryotes, including those of clinical and biotechnological significance. Only a few of the isolates were cryophilic that might be endangered by the rapid loss of the glacier. A number of isolates from the snowpack are, however, the general colonizers that are blown in by the downward and upward wind from the surrounding terrestrial, moraine, savannah ecosystems and the predominant oceanic weather system.

5.1.4 Composition and diversity of the microbial communities

The microbial richness varied across the glacier and foreland samples based on the analyses of the 16S rDNA of the ASV richness at 3% distance. Foreland rhizosphere samples had the highest ASV richness, followed by the soil samples and glacier registered the lowest community richness between the sampling sites. Plant rhizosphere is rich ecosystem with diverse nutritional factors facilitated by the biogeochemical processes from the mycorrhizal associations and saprophytic organisms (Schutte *et al.,* 2009).

Age has got greater effect on the ASV composition along foreland chrono-sequence in both the rhizosphere and soil samples. The 42-year-old rhizosphere and soil samples had the highest number of ASV composition followed by the 23-year-old samples and finally the 12-year-old foreland had the list number of ASV composition. Due to the steady directional non-replacement model (Vater and Matthews, 2013), the rhizosphere samples shared 271 ASVs a cross the different ages along the foreland chrono-sequence. In contrary, 315 ASVs were shared across the ages in the soil samples due to the intermediate directional replacement model at the mid-23-year-old samples. The intermediate model creates intra-species inter-species dynamics as microbial communities tend to re-associate to create a stable niche for the mycorrhizal associations. After the intermediate dynamics, the old foreland experiences a logarithmic increase in the ASV composition due to the increase of the microbial community structures.

Lewis glacier is a unique tropical ecosystem due to the lack of accumulation layer, which is an important region of a typical glacier to hold the glacier content (Prinz *et al.,* 2012). Generally, the supraglacial surfaces are dominated by the phototrophic *Cyanobacteria,* that controls the availability of the inorganic nitrogen and phosphate compounds (Uetake *et al.,* 2010) in the glacier environment. The lack of accumulation layer in Lewis glacier contributes to the rapid loss of the glacier content, which includes the flow of supraglacial microorganisms, snow water and nutrients by the crevassing processes, which couples the supraglacial and subglacial ecosystems through hydrological transfer within the glacier

ecosystem (Hodson *et al.,* 2008). These courses explain why the glacier samples had the lowest ASV composition of 836. 185 ASVs were, however, shared across the rhizosphere, soil and glacier samples, while 380 ASVs were shared between the glacier and soil, 43 ASVs shared between the glacier and rhizosphere and 2436 ASVs shared between the soil and rhizosphere samples. The upper site of the glacier had 680 ASVs compared to the lower 313 ASVs from the lower site of the glacier. The difference in the ASV composition is due to the orientation of the glacier sites. The upper and lower sites are structured into two sites ranging between 4875m and 4675m above sea level (Prinz *et al.,* 2012) with slope gradients of 17.4 degrees and 19.8 degrees, respectively. The slope gradient indicates that the upper site is relatively flat while the lower site is steep.

The upper site therefore has higher retention time to the channelized tillage of the glacier content. The longer the retention time the higher the possible number of ASV amplicons, which are slowly lost to the melt water deposits at the glacier melting point. The steep lower site of the glacier facilitates the deglaciation process through avalanching and crevassing of the glacier content (Hodson *et al.,* 2008). The lower site is also characterized by the flowing subglacial river, that leads the channelized tillage to the foreland endogenous glacier deposits. Crevassing significantly contributes to the rapid loss of the glacier content by enhancing glacier basal sliding and till deformation (Paterson, 1994).

The ASV composition between the glacier and the foreland is due to the endogenous glacier melt deposits and the aerial influxes that colonize the primary foreland and the glacier (Fogg *et al.,* 2008; Scanlon *et al.,* 2015). The competition at the primary foreland affects the ASV composition of the subsequent old ecosystem leading to the lowest, 43 ASVs shared between the glacier and the rhizosphere samples. The horizontal and the vertical dynamics across the foreland chrono-sequence explains the high number of ASVs shared between the soil and rhizosphere samples. Most of the phyla that help in the degradation of the organic compounds in the soil also play a significant role in nitrogen fixation in the mycorrhizal associations in the plant rhizosphere (Hodson *et al.,* 2008).

The rarefaction curve analyses of the 16S rDNA sampling status of the glacier and foreland samples shown that a number of samples were far from being exhaustively analysed. Only eight samples demonstrated the attainment sequencing depth. The rarefied curve demonstrate that ASV amplicon sequences are more diverse in the Lewis glacier and foreland samples. The NMDS Weighted analyses of the 16S rDNA from the glacier samples revealed some ASVs are shared across glacier habitats. The phenomenon was more pronounced in 16S rDNA dataset, indicating that DNA pool contained a ‘‘seed bank’’ of inactive organisms (Jones & Lennon, 2010.

The NMDS analyses were supported by the ASV and taxonomic composition, which divided the datasets into four ellipses based on sampling sites and three ellipses based on the sample type. The significant differences in microbial community structure in the samples at 3% distance level of confidence could be attributed to differences in albedo reflectance on the glacier surfaces, sample type of specific sites during sampling and the age of the foreland samples. The NMDS clusters of this scenario has also been observed from the microbial community compositions of the hot springs in Eritrea (Ghilamicael *et al.,* 2017).

5.1.5 Indicator species in the glacier and its foreland

The indicator species analysis allowed for the identification of the characteristic phyla and subsequent species that were associated with a particular sample types and thus leading to changes in observed patterns. For example ASVs affiliated to phylum *Verrucomicrobia* were associated with the rhizosphere and the soil samples while species that were affiliated to phylum *Actinobacteria* were associated to glacier samples.

This study might have been affected by the sample collection biasness as the analyses was done with the DNA samples, meaning that, only inactive or dormant prokaryotes were targeted. These, however, could have had little influence on ecosystem functioning at the time of sampling. Moreover, the identification of indicator species might not entirely shed light on ecosystem roles related to microbial community structures due to limited

knowledge about the majority of other microorganisms that were not captured by the 16S primer as well as their ecological functions. Nevertheless, a similar study has been done by Hartmann et al. (2012) and Rime et al. (2015). These two studies emphasized that indicator species analyses is best to analyse the role of organisms in a specific ecosystem.

During deglaciation, succession processes include soil development, accumulation of carbon, accumulation of nitrogen and other nutrients along the foreland chrono-sequence (Schütte *et al.,* 2009; Bernasconi *et al.,* 2011; Bajerski & Wagner, 2013). This is due to exogenous colonization by plants (Knelman *et al.,* 2012) and glacial endogenous deposits through the melt water (Li *et al.,* 2014; Scanlon *et al.,* 2015). The process finally leads to higher microbial biomass (Welc *et al.,* 2012) and activity (Göransson *et al.,* 2011). The phenomenon, therefore, results to variation in the taxa functional roles in the glacier and foreland samples. The dynamics in microbial community structure and diversity related to accumulation of organic matter and nutrients during soil development have been reported in Alpine, Arctic and Antarctic glacier forelands (Schütte *et al.,* 2009; Zumsteg *et al.,* 2012; Bajerski & Wagner, 2013).

The association of phyla *Gemmatimonadetes* and *Cyanobacteria* to the soil and glacier ecosystems have been reported elsewhere (Rime *et al.,* 2015). Rime et al. (2015) studied the indicator species of the foreland soil depths. This study was, however, similar to the current study where samples were collected at the foreland shallow soils of 5 cm to 20 cm. the barren soils of Lewis glacier foreland were revealed to be controlled by the phyla *Gemmatimonadetes.* The representatives to this phyla are suggested to have a broad physiological potentials ranging from the glacier foreland (*Zumsteg et al.,* 2012; Bajerski & Wagner, 2013), alpine and arid environments (DeBruyn & Nixon, 2011), agricultural soils (Chaudhry *et al.,* 2012), to the freshwater lakes (Zeng *et al.,* 2014). The representatives to this phyla are ubiquitous due to their varied types of metabolism (Rime *et al.,* 2015), but their ecological roles are poorly understood because they are rarely cultivated (DeBruyn *et al.,* 2013; Zeng *et al.,* 2014). Only one genus has been cultivated (Zhang *et al.,* 2003). It was also not possible to cultivate representatives of this phylum

from Lewis glacier and could only be identified by the next generation sequencing techniques.

The representatives of phylum *Cyanobacteria* colonize glacier ecosystem in abundance and control photosynthetic activities (Hodson *et al.,* 2008). This has been revealed by the indicator species analyses at the phylum level and genus level. Other genera that are affiliated to phylum *Cyanobacterium* have been reported to have mechanisms of surviving while controlling the environmental activities. Genus *Phormidium* has the ability to retain water and produce pigments and sheath protecting them against high UV exposure (Hallenbeck, 2017), which puts high selective pressure on microbial communities (Meola *et al.,* 2014). This mechanisms has enabled species affiliated to *Cyanobacteria* to be isolated from the primary deglaciated foreland (Meola *et al.,* 2014), rocks (Sigler *et al.,* 2003) and desert soil crusts (Garcia-Pichel *et al.,* 2003; Freeman *et al.,* 2009).

Genera like *Tychonema, Phormidesmis* and *Chamaesiphon* that are affiliated to phylum *Cyanobacteria* were revealed to occur as indicators on the glacier with an indicator values of 0.84, 0.74 and 0.68 respectively. Their occurrence in high values corresponds to the ability to retain water and produce pigments that provide sheath protecting fecundity and longevity when carrying out environmental activities. Their functions in the foreland are, however, affected by the fluctuations in water regime and rapid exposure to UV radiations (Rime *et al.,* 2015).

Species affiliated to phylum *Verrucomicrobia* are known to be abundant summer microbial community structures in the soil (Lipson & Schmidt, 2004). They have also been found to degrade plant carbon compounds while thriving under oligotrophic conditions (Janssen *et al.,* 2005), which might be a characteristic of foreland barren soils with sparsely populated vascular plants like Lewis glacier foreland. Dunfield *et al.* (2007) reported that representatives to phylum *Verrucomicrobia* can also oxidize methane in acidic soils. They are also able to colonize the hot springs of the Eritrean lakes (Ghilamicael *et al.,* 2017) in Africa.

The ability of *Verrucomicrobia* affiliates to oxidize methane is an important role to nitrogen cycling in the plant rhizosphere. Elsewhere (Neff *et al.,* 1994), it was reported that, there is a steady decrease in methane oxidation in nitrogen-fertilizer amended soils that do not require microbial function. In addition, phylum *Verrucomicrobia* diversity increase with the increase in soil age (Nemergut *et al.,* 2007). It is a common soil bacteria, which is very difficult to cultivate using the culture-dependent methods. There are other suggested factors such as the mutualistic association with other groups of other microorganisms that may have made it difficult to isolate pure isolates of these species (Nemergut *et al.,* 2007).

The functional role of *Verrucomicrobia* as the indicators to rhizosphere samples is strongly supported by the work of Nemergut et al. (2007), which reported a higher proportion of sequences of such phylum to be related to older soils. Older soils are usually characterized by the diverse vascular plants like the *Senecio keniophytun* as the first pioneer sparsely distributed plants in the newly deglaciated foreland. The *Senecio keniophytun* vegetation is then followed by the high-cover of *Carex monostachya, Agrostis spp., Carduus platyphyllus, Arabis alpina, Senecio keniophytum* and *Lobelia telekii,* which stabilize the foreland soil for mycorrhizal associations (Mizuno & Fujita, 2014). The stable soil also creates a suitable environment for the higher phyla like the *Verrucomicrobia.*

5.1.6 Relationship between microbial community structure on the glacier and environmental variables

The correlation among the major environmental factors on Lewis glacier was exceptionally strong. Cryoconite material and minerals were more correlated indicating their major contribution to microbial community structure on the glacier. Cryoconite material on the glacier surface couples the microbial community structure (Hoham and Duval, 2001; Takeuchi *et al.,* 2001; Segawa *et al.,* 2005; Turchetti *et al.,* 2008; Simon *et al.,* 2009; Hodson *et al.,* 2010; Langford *et al.,* 2010) and complex extracellular polymeric

substances (Hodson *et al.,* 2008) making it act as a micro-biogeochemical reactor (Hodson *et al.,* 2008; Stibal *et al.,* 2010). The polymeric substances also contribute to the inefficient network of pathways and finally joins the channelized large volume of flowing melts (Walder & Fowler, 1994; Hodgkins *et al.,* 2013) to the foreland endogenous glacier deposits.

Lewis glacier surface has got lower reflectance of the sunlight irradiation. The lower rate of wavelength reflection is a factor of consideration to rapid loss of the glacier content (Thomson *et al.,* 2009; Rabatel *et al.,* 2013). The lower site of the glacier has relatively higher reflectance values as compared to the upper site. The difference in the reflectance levels is important to the variation in the ASV composition on the glacier surfaces. The retention of the channelized tillage on the upper site and supraglacial content contributes to the lower reflectance.

Cryoconite material on the glacier surface couples the microbial community structure (Hoham & Duval, 2001; Takeuchi *et al.,* 2001; Segawa *et al.,* 2005; Turchetti *et al.,* 2008; Simon *et al.,* 2009; Hodson *et al.,* 2010; Langford *et al.,* 2010) and complex extracellular polymeric substances (Hodson *et al.,* 2008) making it act as a micro-biogeochemical reactor (Hodson *et al.,* 2008; Stibal *et al.,* 2010). The polymeric substances also contribute to the inefficient network of pathways and finally joins the channelized large volume of flowing melts (Walder & Fowler, 1994; Hodgkins *et al.,* 2013) to the foreland endogenous glacier deposits.

The gradient and slope of the glacier surface is suggested elsewhere (Zarsky *et al.,* 2013) to affect the development of the cryoconite aggregates due to erosive action of the meltwater. The lower site of the glacier has got relatively lower albedo reflectance due to continuous glacier content flow to the subglacial pressure melting points. The flowing melt volume consist of the supraglacial and subglacial ecosystems. The phototrophic organisms that colonize the snowpack and the supraglacial zones such as the snow algae, phytoflagellates and cyanobacteria form part of the melt flow (Hodson *et al.,* 2008). The

steady turbulence (Hodson *et al.,* 2008) within the glacier ecosystem also results to loss of cryoconite materials content, that determine the albedo reflectance levels along the glacier surface. The loss of the dark particles from the supraglacial surface means an improved reflectance surface.

Generally, the microorganisms in the cryoconite holes is much greater than that of the ice cover and the mats (Takeuchi *et al.,* 2000; Porazinska *et al.,* 2004; (Säwström *et al.,* 2002; Hodson *et al.,* 2008). However, the variability of the microbial community structure is dependent on the environmental conditions; especially the glacier slopes and inter-cryoconite hole mixing that leads to continuous colony flushing by the large volume melt water and ecosystem homogenization, respectively (Hodson *et al.,* 2008).

Cryoconite holes from the polar glaciers have been reported to harbour filamentous fungi and yeast (Margesin *et al.,* 2003). However, most cryoconite are known to be dominated by the *Cyanobacteria*, phytoflagallate bacteria, algae and other fungi (Kaštovská *et al.,* 2007; Hodson *et al.,* 2008). Viruses are no exception to the cryoconite regions (Säwström *et al.,* 2002). Viruses have been isolated from the melt water, cryoconite granules and sediments (Mac Donell & Fitzsimons, 2008). The role of viruses on glacier is highly significant as they control the bacterial and fungal availability through cell lysis of the hosts (Hodson *et al.,* 2008). The viral activity, however, requires further investigation to understand their role and impact on the glacier ecosystem.

5.2 Conclusions

1. Lewis glacier has high prokaryotic abundance across the glacier and the foreland chrono-sequence. After taxonomic filter for the above average taxa abundance, 16 known phyla and two unidentified phyla were analysed. The dominant phyla varied from the upper to lower glacier and from the side-foreland to glacier foreland. The minor phyla were, however, distributed almost evenly across the samples.
2. The study revealed four predominant phyla, phylum *Cyanobacteria* which was the most prominent on the glacier sites while phyla *Proteobacteria, Bacteriodetes* and *Actinobacteria* dominated the foreland chrono-sequence.
3. Seventeen isolates from the snowpack were sequenced. Their sequences were closely related to three phyla *Firmicutes, Proteobacteria,* and *Actinobacteria. Firmicutes* were the most abundant on the snowpack.
4. Prokaryotes in Lewis glacier are distributed across the glacier and foreland samples. This phenomena is, however, affected by the sample type, sampling site, altitude and age of the deglaciated foreland soil along the foreland chrono-sequence.
5. The glacier ecosystem changes and species preferences is controlled by the species related to phylum *Cyanobacteria* whereas species related to phyla *Verrucomicrobia* and *Gemmatimonadete* were the indictors for the rhizosphere and soil foreland respectively.
6. The environmental factors significantly affected the distribution of bacterial community structures. the two factor order analyses revealed that organic matter and minerals had a strong correlation to affect the microbial distribution.

5.3 Recommendations

1. There is need to explore the physiological and functional roles of the isolated bacteria on Lewis glacier which is rapidly disappearing.
2. The presence of the nitrogen fixing and biodegrading bacteria on Lewis glacier is exceptionally very important and requires further analyses due to methane oxidation.
3. The abundance of phylum *Firmicutes* on the snowpack indicates an influx of soil particles and bioaerosols on the glacier surface. This aspect is, however, not conclusive and there is need for the analyses for the bioaerosol on Lewis glacier.
4. Indicator species are the most important bacteria from the glacier that determine the changes on the glacier and its foreland. These organisms can be used in future to predict the status of bacterial ecology on the glacier.
5. There is need to explore the significance of organic matter and minerals which were more correlated environmental variables within the cryoconite holes.
6. There is need for a collaboration of experts in the fields of geology, plant ecologists, population geneticists, molecular biologists to help date the glacier succession comprehensively.

REFERENCES

Allali, I., Arnold, J. W., Roach, J., Cadenas, M. B., Butz, N., Hassan, H. M., ... & Azcarate-Peril, M. A. (2017). A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. *BMC microbiology*, *17*(1), 194.

Amato, P., Doyle, S., & Christner, B. C. (2009). Macromolecular synthesis by yeasts under frozen conditions. *Environmental microbiology*, *11*(3), 589-596.

Anesio, A. M., & Laybourn-Parry, J. (2012). Glaciers and ice sheets as a biome. *Trends in Ecology & Evolution*, *27*(4), 219-225.

Anesio, A. M., Hodson, A. J., Fritz, A., Psenner, R., & Sattler, B. (2009). High microbial activity on glaciers: importance to the global carbon cycle. *Global Change Biology*, *15*(4), 955-960.

Anesio, A. M., Lutz, S., Chrismas, N. A., & Benning, L. G. (2017). The microbiome of glaciers and ice sheets. *npj Biofilms and Microbiomes*, *3*(1), 10.

Bajerski, F., & Wagner, D. (2013). Bacterial succession in Antarctic soils of two glacier forefields on Larsemann Hills, East Antarctica. FEMS *Microbiology Ecology*, 85(1), 128-142.

Bardgett, R. D., Richter, A., Bol, R., Garnett, M. H., Bäumler, R., Xu, X., ... & Wanek, W. (2007). Heterotrophic microbial communities use ancient carbon following glacial retreat. *Biology Letters*, *3*(5), 487-490.

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. Nat. 438: 303.

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, *438*(7066), 303-309.

Bell, R. E., Ferraccioli, F., Creyts, T. T., Braaten, D., Corr, H., Das, I., ... & Studinger, M. (2011). Widespread persistent thickening of the East Antarctic Ice Sheet by freezing from the base. *Science*, *331*(6024), 1592-1595.

Berg, G., Marten, P., & Ballin, G. (1996). *Stenotrophomonas maltophilia* in the rhizosphere of oilseed rape—occurrence, characterization and interaction with phytopathogenic fungi. *Microbiology Research 151*(19), 27.

Berg, G., Roskot, N., & Smalla, K. (1999). Genotypic and phenotypic relationships between clinical and environmental isolates of *Stenotrophomonas maltophilia*. *Journal of Clinical Microbiology*. *37*(3594), 3600.

Bernasconi, S. M. (2008). Weathering, soil formation and initial ecosystem evolution on a glacier forefield: a case study from the Damma Glacier, Switzerland. *Mineralogical Magazine*, *72*(1), 19-22.

Bernasconi, S. M., Bauder, A., Bourdon, B., Brunner, I., Bünemann, E., Chris, I., ... & Frossard, E. (2011). Chemical and biological gradients along the Damma glacier soil chronosequence, Switzerland. *Vadose Zone Journal*, *10*(3), 867-883.

Bhattacharya, D., Nagpure, A., & Gupta, R. K. (2007). Bacterial chitinases: properties and potential. *Crit. Rev. Biotech*, *27*(21), 28.

Björkman, M. P., Kühnel, R., Partridge, D. G., Roberts, T. J., Aas, W., Mazzola, M., ... & Isaksson, E. (2013). Nitrate dry deposition in Svalbard. *Tellus B*, *65*.

Blatter, H., & Hutter, K. (1991). Polythermal conditions in Arctic glaciers. *Journal of Glaciology*, *37*(126), 261-269.

Boetius, A., Anesio, A. M., Deming, J. W., Mikucki, J. A., & Rapp, J. Z. (2015). Microbial ecology of the cryosphere: sea ice and glacial habitats. *Nature Reviews Microbiology*, *13*(11), 677.

Bonnet, E., & Van de Peer, Y. (2002). zt: A sofware tool for simple and partial mantel tests. *Journal of Statistical software*, *7*(10), 1.

Bormann, B. T. & Sidle, R. C. (1990). Changes in productivity and distribution of nutrients in a chronosequence at Glacier Bay National Park, Alaska. *Journal of Ecology*, *78*(561), 578.

Bowman, J. P., McCammon, S. A., Brown, M. V., Nichols, D. S., & McMeekin, T. A. (1997). Diversity and association of psychrophilic bacteria in Antarctic sea ice. *Applied Environmental Microbiology*, *63*(3068), 3078.

Brankatschk, R., Töwe, S., Kleineidam, K., Schloter, M., & Zeyer, J. (2011). Abundances and potential activities of nitrogen cycling microbial communities along a chronosequence of a glacier forefield. *The ISME journal*, *5*(6), 1025-1037.

Brizzio, S., Turchetti, B., De Garcia, V., Libkind, D., Buzzini, P., & Van Broock, M. (2007). Extracellular enzymatic activities of basidiomycetous yeasts isolated from glacial and subglacial waters of northwest Patagonia (Argentina). *Canadian Journal of Microbiology*, *53*(4), 519-525.

Butinar, L., Spencer-Martins, I., & Gunde-Cimerman, N. (2007). Yeasts in high Arctic glaciers: the discovery of a new habitat for eukaryotic microorganisms. *Antonie van Leeuwenhoek*, *91*(3), 277-289.

Buzzini, P., Branda, E., Goretti, M., & Turchetti, B. (2012). Psychrophilic yeasts from worldwide glacial habitats: diversity, adaptation strategies and biotechnological potential. *FEMS microbiology ecology*, *82*(2), 217-241.

Cameron, K. A., Hodson, A. J., & Osborn, A. M. (2012). Structure and diversity of bacterial, eukaryotic and archaeal communities in glacial cryoconite holes from the Arctic and the Antarctic. *FEMS microbiology ecology*, *82*(2), 254-267.

Chapin, F. S., Walker, L. R., Fastie, C. L. & Sharman, L. C. (1994). Mechanisms of primary succession following deglaciation at Glacier Bay, Alaska. *Ecology Monogr*, *64*(149), 175.

Chaudhry, V., Rehman, A., Mishra, A., Chauhan, P. S., & Nautiyal, C. S. (2012). Changes in bacterial community structure of agricultural land due to long-term organic and chemical amendments. *Microbial ecology*, 64(2), 450-460.

Christner, B. C., Kvitko II, B. H., & Reeve, J. N. (2003). Molecular identification of bacteria and eukarya inhabiting an Antarctic cryoconite hole. *Extremophiles*, *7*(3), 177-183.

Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., & Tiedje, J. M. (2013). Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 42: D633-D642.

Crocker, R. L. & Major, J. (1955). Soil development in relation to vegetation and surface age at Glacier Bay, Alaska. *Journal of Ecology*, *43*(427), 448.

Cuffey, K. M., & Paterson, W. S. B. (2010). *The physics of glaciers*. Academic Press.

Cullen, N. J., Mölg, T., Kaser, G., Hussein, K., Steffen, K., & Hardy, D. R. (2006). Kilimanjaro Glaciers: Recent areal extent from satellite data and new

interpretation of observed 20th century retreat rates. *Geophysical Research Letters*, *33*(16).

Davey, M., Blaalid, R., Vik, U., Carlsen, T., Kauserud, H., & Eidesen, P. B. (2015). Primary succession of Bistorta vivipara (L.) Delabre (Polygonaceae) root-associated fungi mirrors plant succession in two glacial chronosequences. *Environmental Microbiology, 17*(2777), 2790.

De García, V., Brizzio, S., Libkind, D., Buzzini, P., & Van Broock, M. (2007). Biodiversity of cold-adapted yeasts from glacial meltwater rivers in Patagonia, Argentina. *FEMS microbiology ecology*, *59*(2), 331-341.

De Smet, W. H., & Van Rompu, E. A. (1994). Rotifera and Tardigrada from some cryoconite holes on a Spitsbergen (Svalbard) glacier. *Belgian Journal of Zoology*, *124*(27), 27.

DeBruyn, J. M., Nixon, L. T., Fawaz, M. N., Johnson, A. M., & Radosevich, M. (2011). Global biogeography and quantitative seasonal dynamics of Gemmatimonadetes in soil. Applied and environmental microbiology, AEM-05005.

Deming, J. W. (2002). Psychrophiles and polar regions. *Current Opinions in Microbiology, 5*(301), 309.

Downie, C. (1964). Glaciations of Mount Kilimanjaro, Northeast Tanganyika. *Geological Society of America Bulletin*, *75*(1), 1-16.

Duane, W. J., Pepin, N. C., Losleben, M. L., & Hardy, D. R. (2008). General characteristics of temperature and humidity variability on Kilimanjaro, Tanzania. *Arctic, Antarctic, and Alpine Research*, *40*(2), 323-334.

Dulic, T., Meriluoto, J., Maleševic, T. P., Gajic, V., Vazic, T., Tokodi, N., ... & Svircev, Z. (2016). Cyanobacterial diversity and toxicity of biocrusts from the Caspian Lowland loess deposits, North Iran. *Quaternary International*, *30*(1), 12.

Dunfield, P. F., Yuryev, A., Senin, P., Smirnova, A. V., Stott, M. B., Hou, S., ... & Wang, J. (2007). Methane oxidation by an extremely acidophilic bacterium of the phylum Verrucomicrobia. *Nature, 450*(7171), 879.

Edwards, A., Pachebat, J. A., Swain, M., Hegarty, M., Hodson, A. J., Irvine-Fynn, T. D., ... & Sattler, B. (2013). A metagenomic snapshot of taxonomic and functional diversity in an alpine glacier cryoconite ecosystem. *Environmental Research Letters*, *8*(3), 035003.

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

Fogg, G. E., Convey, P., Fritsen, C. H., Gili, J. M., Gradinger, R., Laybourn-Parry, J., ... & Walton, D. W. H. (2008). *The biology of polar regions*. Oxford University Press.

Foreman, C. M., Sattler, B., Mikucki, J. A., Porazinska, D. L., & Priscu, J. C. (2007). Metabolic activity and diversity of cryoconites in the Taylor Valley, Antarctica. *Journal of Geophysical Research: Biogeosciences*, *112*(G4).

Freeman, K. R., Pescador, M. Y., Reed, S. C., Costello, E. K., Robeson, M. S., & Schmidt, S. K. (2009). Soil CO2 flux and photoautotrophic community composition in high-elevation,'barren'soil. *Environmental microbiology*, *11*(3), 674-686.

Garcia-Pichel, F., Johnson, S. L., Youngkin, D., & Belnap, J. (2003). Small-scale vertical distribution of bacterial biomass and diversity in biological soil crusts from arid lands in the Colorado Plateau. *Microbial Ecology*, *46*(3), 312-321.

Gardner, A. S., Moholdt, G., Wouters, B., Wolken, G. J., Burgess, D. O., Sharp, M. J., ... & Labine, C. (2011). Sharply increased mass loss from glaciers and ice caps in the Canadian Arctic Archipelago. *Nature*, *473*(7347), 357-360.

Ghilamicael, A. M., Boga, H. I., Anami, S. E., Mehari, T., & Budambula, N. (2017). Physical and Chemical Characteristics of Five Hot Springs in Eritrea.Jones and Lennon, 2010

Ghilamicael, A. M., Budambula, N. L., Anami, S. E., Mehari, T., & Boga, H. I. (2018). Thermotolerant bacteria of biotechnological potential from hot springs in Eritrea. *African Journal of Microbiology Research*, *12*(22), 512-524.

Göransson, H., Venterink, H. O., & Bååth, E. (2011). Soil bacterial growth and nutrient limitation along a chronosequence from a glacier forefield. *Soil Biology and Biochemistry*, *43*(6), 1333-1340.

Grebmeier, J. M., & Barry, J. P. (1991). The influence of oceanographic processes on pelagic-benthic coupling in polar regions: a benthic perspective. *Journal of Marine Systems*, *2*(3), 495-518.

Hallbeck, L. (2009). Microbial processes in glaciers and permafrost. A literature study on microbiology affecting groundwater at ice sheet melting, SKB Rapport R-09-37.

Hallenbeck, P. C. (2017). Desert Crusts. In Modern Topics in the Phototrophic Prokaryotes (217-242).

Harding, T., Jungblut, A. D., Lovejoy, C., & Vincent, W. F. (2011). Microbes in high arctic snow and implications for the cold biosphere. *Applied Environmental Microbiology*, *77*(3234), 3243.

Hartmann, M., Howes, C. G., VanInsberghe, D., Yu, H., Bachar, D., Christen, R., ... & Mohn, W. W. (2012). Significant and persistent impact of timber harvesting on soil microbial communities in Northern coniferous forests. *The ISME Journal, 6*(12), 2199.

Hastenrath, S. (1983). Net balance, surface lowering, and ice-flow pattern in the interior of Lewis Glacier, Mount Kenya, Kenya. *Journal of Glaciology*, *29*(392), 402.

Hastenrath, S. (2005). The Glaciers of Mount Kenya 1899–2004 (Veränderungen der Gletscher am Mount Kenya 1899–2004). *Erdkunde*, 120-125.

Hastenrath, S. (2006). Diagnosing the decaying glaciers of equatorial East Africa. *Meteorologische Zeitschrift*, *15*(3), 265-271.

Hastenrath, S. (2008). *Recession of equatorial glaciers: a photo documentation* (Vol. 25). Madison, WI, US: Sundog Publishing.

Hastenrath, S. (2010). Climatic forcing of glacier thinning on the mountains of equatorial East Africa. *A Journal of the Royal Meteorological Society*, *30*(1), 146-152.

Hock, R., & Holmgren, B. (2005). A distributed surface energy-balance model for complex topography and its application to Storglaciären, Sweden. *Journal of Glaciology*, *51*(172), 25-36.

Hodson, A., Anesio, A. M., Tranter, M., Fountain, A., Osborn, M., Priscu, J., ... & Sattler, B. (2008). Glacial ecosystems. *Ecological monographs*, *78*(1), 41-67.

Hodson, A., Kohler, J., Brinkhaus, M., & Wynn, P. (2005). Multi-year water and surface energy budget of a high-latitude polythermal glacier: evidence for overwinter water storage in a dynamic subglacial reservoir. *Annals of Glaciology*, *42*(42), 46.

Hoham, R. W., & Duval, B. (2001). Microbial ecology of snow and freshwater ice with emphasis on snow algae. *Snow ecology: an interdisciplinary examination of snow-covered ecosystems. Cambridge University Press, Cambridge*, 168-228.

Irvine-Fynn, T. D., Hodson, A. J., Moorman, B. J., Vatne, G., & Hubbard, A. L. (2011). Polythermal glacier hydrology: A review. *Reviews of Geophysics*, *49*(4).

Johnsen, A. R., Wick, L. Y., & Harms, H. (2005). Principles of microbial PAH-degradation in soil. Environmental pollution, 133(1), 71-84.

Jones, H. G., Pomeroy, J. W., Davies, T. D., Tranter, M., & Marsh, P. (1999). in Arctic snow cover: landscape form, in-pack gas concentration gradients, and the implications for the estimation of gaseous fluxes. *Hydrology Process*, *13*(2977), 2989.

Jumpponen, A. (2003). Soil fungal community assembly in a primary successional glacier forefront ecosystem as inferred from rDNA sequence analyses. *New Phytologist 158*(569), 578.

Jungblut, A. D., Vincent, W. F., & Lovejoy, C. (2012). Eukaryotes in Arctic and Antarctic cyanobacterial mats. *FEMS Microbiology Ecology*, *82*(2), 416-428.

Kaser, G. (1999). A review of the modern fluctuations of tropical glaciers. *Global Planet Change*, *22*(93), 103.

Kaštovská, K., Elster, J., Stibal, M., & Šantrůčková, H. (2005). Microbial assemblages in soil microbial succession after glacial retreat in Svalbard (High Arctic). *Microbial Ecology*, *50*(3), 396-407.

Kekonen, T., Moore, J., Perämäki, P., Mulvaney, R., Isaksson, E., Pohjola, V., & van de Wal, R. S. (2005). The 800-year-long ion record from the Lomonosovfonna (Svalbard) ice core. *Journal of Geophysical Research: Atmospheres*, *110*(D7).

Klein, A., & Kincaid, J. (2007). A Reassessment of the Satellite Record of Glacier Change in the Rwenzori Mountains, East Africa. In *64th Eastern Snow Conference, St. John's, Newfoundland, Canada*.

Knelman, J. E., Legg, T. M., O'Neill, S. P., Washenberger, C. L., González, A., Cleveland, C. C., & Nemergut, D. R. (2012). Bacterial community structure and function change in association with colonizer plants during early primary succession in a glacier forefield. *Soil Biology and Biochemistry*, *46*(172), 180.

Kohler, J., James, T. D., Murray, T., Nuth, C., Brandt, O., Barrand, N. E., ... & Luckman, A. (2007). Acceleration in thinning rate on western Svalbard glaciers. *Geophysical Research Letters*, *34*(18).

Kol, E. (1969). The red snow of Greenland. II. *Acta Bot Acad Sci Hung*, *15*(3-4), 281-289.

Kruss, P. (1983). Climate change in East Africa: A numerical simulation from the 100 years of terminus record at Lewis Glacier, Mount Kenya. *Zeitschrift für Gletscherkunde und Glazialgeologie*, *19*(1), 43-60.

Kühnel, R., Roberts, T. J., Björkman, M. P., Isaksson, E., Aas, W., Holmén, K., & Ström, J. (2012). 20-year climatology of and wet deposition at Ny-Ålesund, Svalbard. *Advances in Meteorology*, *2011*.

Kustas, W. P., Rango, A., & Uijlenhoet, R. (1994). A simple energy budget algorithm for the snowmelt runoff model. *Water Resources Research*, *30*(5), 1515-1527.

Langford, H., Hodson, A., Banwart, S., & Bøggild, C. (2010). The microstructure and biogeochemistry of Arctic cryoconite granules. *Annals of Glaciology*, *51*(56), 87-94.

Lennon, J. T., & Jones, S. E. (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nature reviews microbiology*, *9*(2), 119.

Li, G., Li, P., Liu, Y., Qiao, L., Ma, Y., Xu, J., & Yang, Z. (2014). Sedimentary system response to the global sea level change in the East China Seas since the last glacial maximum. *Earth-Science Reviews*, *139*(390), 405.

Lipson, D. A., & Schmidt, S. K. (2004). Seasonal changes in an alpine soil bacterial community in the Colorado Rocky Mountains. *Applied and environmental microbiology*, *70*(5), 2867-2879.

MacDonell, S., & Fitzsimons, S. (2008). The formation and hydrological significance of cryoconite holes. *Progress in Physical Geography*, *32*(6), 595-610.

Makhalanyane, T. P., Valverde, A., Velázquez, D., Gunnigle, E., Van Goethem, M. W., Quesada, A., & Cowan, D. A. (2015). Ecology and biogeochemistry of cyanobacteria in soils, permafrost, aquatic and cryptic polar habitats. *Biodiversity and Conservation*, *24*(4), 819-840.

Margesin, R., Gander, S., Zacke, G., Gounot, A. M., & Schinner, F. (2003). Hydrocarbon degradation and enzyme activities of cold-adapted bacteria and yeasts. *Extremophiles*, *7*(6), 451-458.

Margesin, R., Hämmerle, M., & Tscherko, D. (2007). Microbial activity and community composition during bioremediation of diesel-oil-contaminated soil: effects of hydrocarbon concentration, fertilizers, and incubation time. *Microbial Ecology*, *53*(2), 259-269.

Matthews, J. A. (1992). The ecology of recently-deglaciated terrain: a geoecological approach to Glacier Forelands and primary succession. Cambridge, UK: Cambridge University Press.

McCarroll, D. (2000). Spectral roughness of glaciated bedrock geomorphic surfaces: implications for glacier sliding. *Journal of Geophysical Research*, *105*(B9), 21-295.

McConnell, J. R., Edwards, R., Kok, G. L., Flanner, M. G., Zender, C. S., Saltzman, E. S., ... & Kahl, J. D. (2007). 20th-century industrial black carbon emissions altered arctic climate forcing. *Science*, *317*(5843), 1381-1384.

Meola, M., Lazzaro, A., & Zeyer, J. (2014). Diversity, resistance and resilience of the bacterial communities at two alpine glacier forefields after a reciprocal soil transplantation. Environmental microbiology, *16*(6), 1918-1934.

Miteva, V., Teacher, C., Sowers, T., & Brenchley, J. (2009). Comparison of the microbial diversity at different depths of the GISP2 Greenland ice core in relationship to deposition climates. *Environmental microbiology*, *11*(3), 640-656.

Mölg, T., Cullen, N. J., Hardy, D. R., Winkler, M., & Kaser, G. (2009). Quantifying climate change in the tropical midtroposphere over East Africa from glacier shrinkage on Kilimanjaro. *Journal of Climate*, *22*(15), 4162-4181.

Mölg, T., Hardy, D. R., & Kaser, G. (2003). Solar-radiation-maintained glacier recession on Kilimanjaro drawn from combined ice-radiation geometry modeling. *Journal of Geophysical Research: Atmospheres*, *108*(D23).

Moreno, P. I., Jacobson, G. L., Lowell, T. V., & Denton, G. H. (2001). Interhemispheric climate links revealed by a late-glacial cooling episode in southern Chile. *Nature*, *409*(6822), 804-808.

Mote, P. W., & Kaser, G. (2007). The shrinking glaciers of Kilimanjaro: can global warming be blamed?. *American Scientist*, *95*(4), 318-325.

Mur, R., Skulberg, O. M., & Utkilen, H. (1999). CYANOBACTERIA IN THE ENVIRONMENT.

Nadeem, S. M., Naveed, M., Zahir, Z. A., & Asghar, H. N. (2013). Plant–microbe interactions for sustainable agriculture: fundamentals and recent advances. In Plant Microbe Symbiosis: Fundamentals and Advances (51-103).

Neff, J. C., Bowman, W. D., Holland, E. A., Fisk, M. C., & Schmidt, S. K. (1994). Fluxes of nitrous oxide and methane from nitrogen-amended soils in a Colorado alpine ecosystem. *Biogeochemistry*, *27*(1), 23-33.

Neff, J. C., Townsend, A. R., Gleixner, G., Lehman, S. J., Turnbull, J., & Bowman, W. D. (2002). Variable effects of nitrogen additions on the stability and turnover of soil carbon. *Nature*, *419*(6910), 915.

Nicol, G. W., Tscherko, D., Embley, T. M. & Prosser, J. I. (2005). Primary succession of soil Crenarchaeota across a receding glacier foreland. *Environmental Microbiology, 7*(337), 347.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., ... & Wagner, H. (2011). vegan: Community ecology package. *R package version*, 117-118.

Palmisano, A. C., & Sullivan, C. W. (1983). Sea ice microbial communities (SIMCO). *Polar Biology*, *2*(3), 171-177.

Parizek, B. R., Alley, R. B., Anandakrishnan, S., & Conway, H. (2002). Sub-catchment melt and long-term stability of ice stream D, West Antarctica. *Geophysical Research Letters*, *29*(8).

Porazinska, D. L., Fountain, A. G., Nylen, T. H., Tranter, M., Virginia, R. A., & Wall, D. H. (2004). The biodiversity and biogeochemistry of cryoconite holes from McMurdo Dry Valley glaciers, Antarctica. *Arctic, Antarctic, and Alpine Research*, *36*(1), 84-91.

Prinz, R., Fischer, A., Nicholson, L., & Kaser, G. (2011). Seventy-six years of mean mass balance rates derived from recent and re-evaluated ice volume measurements on tropical Lewis Glacier, Mount Kenya. *Geophysical Research Letters*, *38*(20).

Rabatel, A., Francou, B., Soruco, A., Gomez, J., Cáceres, B., Ceballos, J. L., ... & Scheel, M. (2013). Current state of glaciers in the tropical Andes: a multi-century perspective on glacier evolution and climate change. *The Cryosphere*, *7*(1), 81-102.

Reiners, W. A., Worley, I. A. & Lawrence, D. B. (1971). Plant diversity in a chronosequence at Glacier Bay, Alaska. *Ecology* 62, 376–386.

Rime, T., Hartmann, M., Brunner, I., Widmer, F., Zeyer, J., & Frey, B. (2015). Vertical distribution of the soil microbiota along a successional gradient in a glacier forefield. *Molecular ecology, 24*(5), 1091-1108.

Ryan, R. P., Germaine, K., Franks, A., Ryan, D. J., & Dowling, D. N. (2008). Bacterial endophytes: recent developments and applications. *FEMS Microbiology Letters, 278*(1), 9.

Sanyal, A., Antony, R., Samui, G., & Thamban, M. (2018). Microbial communities and their potential for degradation of dissolved organic carbon in cryoconite hole environments of Himalaya and Antarctica. *Microbiological research*, *208*(32), 42.

Säwström, C., Granéli, W., Laybourn-Parry, J., & Anesio, A. M. (2007). High viral infection rates in Antarctic and Arctic bacterioplankton. *Environmental Microbiology*, *9*(1), 250-255.

Säwström, C., Mumford, P., Marshall, W., Hodson, A., & Laybourn-Parry, J. (2002). The microbial communities and primary productivity of cryoconite holes in an Arctic glacier (Svalbard 79 N). *Polar Biology*, *25*(8), 591-596.

Scanlon, K. E., Head, J. W., & Marchant, D. R. (2015). Volcanism-induced, local wet-based glacial conditions recorded in the Late Amazonian Arsia Mons tropical mountain glacier deposits. *Icarus*, *250*(18), 31.

Schmidt, S. K., & Lipson, D. A. (2004). Microbial growth under the snow: implications for nutrient and allelochemical availability in temperate soils. *Plant and Soil*, *259*(1-2), 1-7.

Schütte, U. M., Abdo, Z., Bent, S. J., Williams, C. J., Schneider, G. M., Solheim, B., & Forney, L. J. (2009). Bacterial succession in a glacier foreland of the High Arctic. *ISME Journal*, *3*(12), 58.

Segawa, T., Miyamoto, K., Ushida, K., Agata, K., Okada, N., & Kohshima, S. (2005). Seasonal change in bacterial flora and biomass in mountain snow from the Tateyama Mountains, Japan, analyzed by 16S rRNA gene sequencing and real-time PCR. *Applied and Environmental Microbiology*, *71*(1), 123-130.

Shain, D. H., Mason, T. A., Farrell, A. H., & Michalewicz, L. A. (2001). Distribution and behavior of ice worms (Mesenchytraeus solifugus) in south-central Alaska. *Canadian Journal of Zoology*, *79*(10), 1813-1821.

Sigler, W. V., & Zeyer, J. (2002). Microbial diversity and activity along the forefields of two receding glaciers. *Microbial Ecology*, *43*(4), 397-407.

Sigler, W. V., Bachofen, R., & Zeyer, J. (2003). Molecular characterization of endolithic cyanobacteria inhabiting exposed dolomite in central Switzerland. *Environmental microbiology, 5*(7), 618-627.

Sigler, W. V., Crivii, S., & Zeyer, J. (2002). Bacterial succession in glacial forefield soils characterized by community structure, activity and opportunistic growth dynamics. *Microbial Ecology*, *44*(4), 306-316.

Simon, C., Wiezer, A., Strittmatter, A. W., & Daniel, R. (2009). Phylogenetic diversity and metabolic potential revealed in a glacier ice metagenome. *Applied and Environmental Microbiology*, *75*(23), 7519-7526.

Singh, P., Singh, S. M., & Dhakephalkar, P. (2014). Diversity, cold active enzymes and adaptation strategies of bacteria inhabiting glacier cryoconite holes of High Arctic. *Extremophiles*, *18*(2), 229-242.

Skidmore, M. L., Foght, J. M., & Sharp, M. J. (2000). Microbial life beneath a high Arctic glacier. *Applied and Environmental Microbiology*, *66*(8), 3214-3220.

Skidmore, M., Anderson, S. P., Sharp, M., Foght, J., & Lanoil, B. D. (2005). Comparison of microbial community compositions of two subglacial environments reveals a possible role for microbes in chemical weathering processes. *Applied and Environmental Microbiology*, *71*(11), 6986-6997.

Skidmore, M., Tranter, M., Tulaczyk, S., & Lanoil, B. (2010). Hydrochemistry of ice stream beds—Evaporitic or microbial effects? *Hydrology Process, 24*(517), 523.

Slater, J. F., Currie, L. A., Dibb, J. E., & Benner, B. A. (2002). Distinguishing the relative contribution of fossil fuel and biomass combustion aerosols deposited at

Summit, Greenland through isotopic and molecular characterization of insoluble carbon. *Atmospheric Environment*, *36*(28), 4463-4477.

Stanier, R. Y., & Bazine, G. C. (1977). Phototrophic prokaryotes: the cyanobacteria. *Annal Rev. Microbiology, 31*(225), 274.

Stibal, M., Šabacká, M., & Kaštovská, K. (2006). Microbial communities on glacier surfaces in Svalbard: impact of physical and chemical properties on abundance and structure of cyanobacteria and algae. *Microbial Ecology*, *52*(4), 644.

Stibal, M., Šabacká, M., & Žárský, J. (2012). Biological processes on glacier and ice sheet surfaces. *Nature Geoscience*, *5*(11), 771-774.

Suzuki, K. I., Sasaki, J., Uramoto, M., Nakase, T., & Komagata, K. (1997). *Cryobacterium psychrophilum* gen. nov., sp. nov., nom. rev., comb. nov., an obligately psychrophilic actinomycete to accommodate "*Curtobacterium psychrophilum*" Inoue and Komagata 1976. *International Journal of Systemic Evolution Microbiology, 47*(474), 478.

Takeuchi, N., Kohshima, S., & Seko, K. (2001). Structure, formation, and darkening process of albedo-reducing material (cryoconite) on a Himalayan glacier: a granular algal mat growing on the glacier. *Arctic, Antarctic, and Alpine Research*, 115-122.

Takeuchi, N., Nishiyama, H., & Li, Z. (2010). Structure and formation process of cryoconite granules on Ürümqi glacier No. 1, Tien Shan, China. *Annals of Glaciology*, *51*(56), 9-14.

Tamura, K., Nei, M., & Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. U.S.A. 101*, 11030-11035.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution, 28*(2731), 2739.

Team, R. C. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2012) ISBN 3-900051-07-0. *www. R-project. org*.

Telling, J., Anesio, A. M., Hawkings, J., Tranter, M., Wadham, J. L., Hodson, A. J., ... & Yallop, M. L. (2010). Measuring rates of gross photosynthesis and net community production in cryoconite holes: a comparison of field methods. *Annals of Glaciology*, *51*(56), 153-162.

Telling, J., Anesio, A. M., Tranter, M., Stibal, M., Hawkings, J., Irvine-Fynn, T., ... & Wadham, J. (2012). Controls on the autochthonous production and respiration of organic matter in cryoconite holes on high Arctic glaciers. *Journal of Geophysical Research: Biogeosciences*, *117*(G1).

ter Braak, C. J., & Schaffers, A. P. (2004). Co-correspondence analysis: a new ordination method to relate two community compositions. *Ecology*, *85*(3), 834-846.

Thompson, L. G., Brecher, H. H., Mosley-Thompson, E., Hardy, D. R., & Mark, B. G. (2009). Glacier loss on Kilimanjaro continues unabated. *Proceedings of the National Academy of Sciences*, *106*(47), 19770-19775.

Tscherko, D., Rustemeier, J., Richter, A., Wanek, W. & Kandeler, E. (2003). Functional diversity of the soil microflora in primary succession across two glacier forelands in the Central Alps. *European Journal of Soil and Science, 54*(685), 696.

Tscherko, D., Rustemeier, J., Richter, A., Wanek, W., & Kandeler, E. (2003). Functional diversity of the soil microflora in primary succession across two glacier forelands in the Central Alps. *European Journal of Soil Science*, *54*(4), 685-696.

Tung, H. C., Bramall, N. E., & Price, P. B. (2005). Microbial origin of excess methane in glacial ice and implications for life on Mars. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(51), 18292-18296.

Turchetti, B., Buzzini, P., Goretti, M., Branda, E., Diolaiuti, G., D'Agata, C., & Vaughan-Martini, A. (2007). Psychrophilic yeasts in glacial environments of Alpine glaciers. *FEMS Microbiology and Ecology, 63*(73), 83.

Turchetti, B., Buzzini, P., Goretti, M., Branda, E., Diolaiuti, G., D'Agata, C., ... & Vaughan-Martini, A. (2008). Psychrophilic yeasts in glacial environments of Alpine glaciers. *FEMS microbiology ecology*, *63*(1), 73-83.

Uetake, J., Kohshima, S., Nakazawa, F., Suzuki, K., Kohno, M., Kameda, T., ... & Fujii, Y. (2006). Biological ice-core analysis of Sofiyskiy glacier in the Russian Altai. *Annals of Glaciology*, *43*(1), 70-78.

Uetake, J., Kohshima, S., Nakazawa, F., Takeuchi, N., Fujita, K., Miyake, T., ... & Nakawo, M. (2011). Evidence for propagation of cold-adapted yeast in an ice core from a Siberian Altai glacier. *Journal of Geophysical Research: Biogeosciences*, *116*(G1).

Uetake, J., Naganuma, T., Hebsgaard, M. B., Kanda, H., & Kohshima, S. (2010). Communities of algae and cyanobacteria on glaciers in west Greenland. *Polar Science*, *4*(1), 71-80.

UNEP. (2005). One Planet, Many People: Atlas of Environmental Change. Nairobi: United Nations Environment Programme.

Van de Wal, R. S. W., Boot, W., Van den Broeke, M. R., Smeets, C. J. P. P., Reijmer, C. H., Donker, J. J. A., & Oerlemans, J. (2008). Large and rapid melt-induced velocity changes in the ablation zone of the Greenland ice sheet. *science*, *321*(5885), 111-113.

Vincent, W. F. (2000). Evolutionary origins of Antarctic microbiota: invasion, selection and endemism. *Antarctic Science*, *12*(03), 374-385.

Vincent, W. F., Whyte, L. G., Lovejoy, C., Greer, C. W., Laurion, I., Suttle, C. A., ... & Mueller, D. R. (2009). Arctic microbial ecosystems and impacts of extreme warming during the International Polar Year. *Polar Science*, *3*(3), 171-180.

Vuille, M., Francou, B., Wagnon, P., Juen, I., Kaser, G., Mark, B. G., & Bradley, R. S. (2008). Climate change and tropical Andean glaciers: Past, present and future. *Earth-science reviews*, *89*(3-4), 79-96.

Wadham, J. L., Bottrell, S., Tranter, M., & Raiswell, R. (2004). Stable isotope evidence for microbial sulphate reduction at the bed of a polythermal high Arctic glacier. *Earth and Planetary Science Letters*, *219*(3), 341-355.

Walker, L. R. & del Moral, R. (2003). Primary succession and ecosystem rehabilitation. Cambridge UK: Cambridge University Press.

Welc, M., Bünemann, E. K., Fließbach, A., Frossard, E., & Jansa, J. (2012). Soil bacterial and fungal communities along a soil chronosequence assessed by fatty acid profiling. *Soil Biology and Biochemistry*, *49*(184), 192.

Wharton Jr, R. A., McKay, C. P., Simmons Jr, G. M., & Parker, B. C. (1985). Cryoconite holes on glaciers. *BioScience*, 499-503.

WSB Paterson. (1994). *The physics of glaciers*. Butterworth-Heinemann.

Yoshitake, S., Uchida, M., Koizumi, H., Kanda, H., & Nakatsubo, T. (2010). Production of biological soil crusts in the early stage of primary succession on a High Arctic glacier foreland. *New Phytol*, *186*, 451-460.

Young, J. A., & Hastenrath, S. (1991). Satellite image atlas of glaciers of the world. *US Geological Survey Professional Paper*, 49.

Zarsky, J. D., Stibal, M., Hodson, A., Sattler, B., Schostag, M., Hansen, L. H., ... & Psenner, R. (2013). Large cryoconite aggregates on a Svalbard glacier support a diverse microbial community including ammonia-oxidizing archaea. *Environmental Research Letters*, *8*(3), 035044.

Zeng, Y., Feng, F., Medová, H., Dean, J., & Koblížek, M. (2014). Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes. *Proceedings of the National Academy of Sciences of the United States of America, 106*(47), 19770-19775.

Zhang, D. C., Wang, H. X., Cui, H. L., Yang, Y., Liu, H. C., Dong, X. Z., & Zhou, P. J. (2007). *Cryobacterium psychrotolerans* sp. nov., a novel psychrotolerant bacterium isolated from the China No. 1 glacier. *International journal of systematic and evolutionary microbiology, 57*(866), 869.

Zhang, H., Sekiguchi, Y., Hanada, S., Hugenholtz, P., Kim, H., Kamagata, Y., & Nakamura, K. (2003). *Gemmatimonas aurantiaca* gen. nov., sp. nov., a Gram-negative, aerobic, polyphosphate-accumulating micro-organism, the first cultured representative of the new bacterial phylum *Gemmatimonadetes* phyl. nov. *International journal of systematic and evolutionary microbiology, 53*(4), 1155-1163.

Zumsteg, A., Luster, J., Göransson, H., Smittenberg, R. H., Brunner, I., Bernasconi, S. M., ... & Frey, B. (2012). Bacterial, archaeal and fungal succession in the forefield of a receding glacier. *Microbial ecology*, *63*(3), 552-564.

APPENDICES

Appendix I: DNA Extraction Reagents

Solution 1

50 mM Tris pH 8.5, 50 mM EDTA pH 8.0

25 % Sucrose solution

Solution 2

10 mM Tris pH 8.5

5 mM EDTA pH 8.0

1 % SDS

Lysozyme 20 mg/ml

RNase A 20 mg/ml

Proteinase K 20 mg/ml

Phenol

Chloroform

Absolute ethanol.

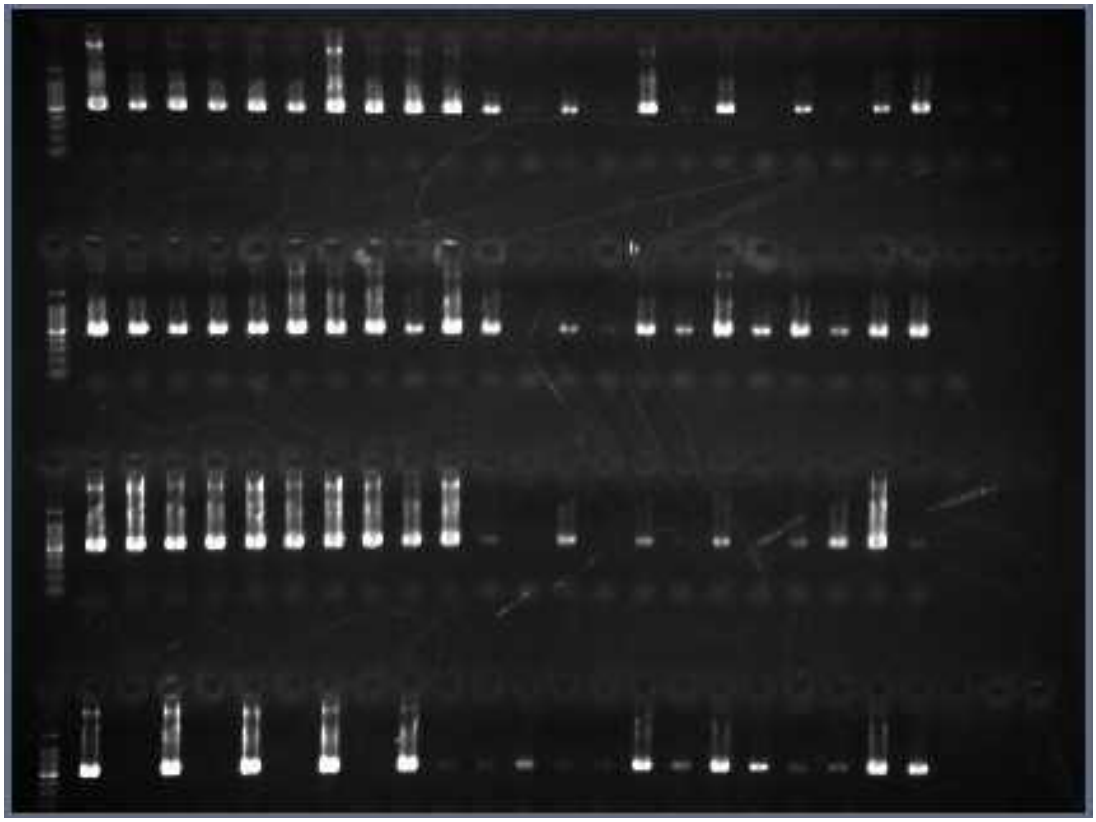
3 M NaCl

Isopropanol

Appendix II: Metadata for the sample analyses

Sample	Nucleic acid	Region	Site	Description	Sample_type	Altitude (masl)	Age (Year)
GF1Rh-1	DNA	Kenya2016	GF1Rh	GF1Rh-1	Rhizosphere	4675	2004
GF1Rh-2	DNA	Kenya2016	GF1Rh	GF1Rh-2	Rhizosphere	4675	2004
GF1Rh-3	DNA	Kenya2016	GF1Rh	GF1Rh-3	Rhizosphere	4675	2004
GF1Rh-4	DNA	Kenya2016	GF1Rh	GF1Rh-4	Rhizosphere	4675	2004
GF1Rh-5	DNA	Kenya2016	GF1Rh	GF1Rh-5	Rhizosphere	4675	2004
GF1S-1	DNA	Kenya2016	GF1S	GF1S-1	Soil	4675	2004
GF1S-2	DNA	Kenya2016	GF1S	GF1S-2	Soil	4675	2004
GF1S-3	DNA	Kenya2016	GF1S	GF1S-3	Soil	4675	2004
GF1S-4	DNA	Kenya2016	GF1S	GF1S-4	Soil	4675	2004
GF1S-5	DNA	Kenya2016	GF1S	GF1S-5	Soil	4675	2004
GF2Rh-1	DNA	Kenya2016	GF2Rh	GF2Rh-1	Rhizosphere	4600	1993
GF2Rh-2	DNA	Kenya2016	GF2Rh	GF2Rh-2	Rhizosphere	4600	1993
GF2Rh-3	DNA	Kenya2016	GF2Rh	GF2Rh-3	Rhizosphere	4600	1993
GF2Rh-4	DNA	Kenya2016	GF2Rh	GF2Rh-4	Rhizosphere	4600	1993
GF2Rh-5	DNA	Kenya2016	GF2Rh	GF2Rh-5	Rhizosphere	4600	1993
GF2S-1	DNA	Kenya2016	GF2S	GF2S-1	Soil	4600	1993
GF2S-2	DNA	Kenya2016	GF2S	GF2S-2	Soil	4600	1993
GF2S-3	DNA	Kenya2016	GF2S	GF2S-3	Soil	4600	1993
GF2S-4	DNA	Kenya2016	GF2S	GF2S-4	Soil	4600	1993
GF2S-5	DNA	Kenya2016	GF2S	GF2S-5	Soil	4600	1993
GF3Rh-1	DNA	Kenya2016	GF3Rh	GF3Rh-1	Rhizosphere	4575	1974
GF3Rh-2	DNA	Kenya2016	GF3Rh	GF3Rh-2	Rhizosphere	4575	1974
GF3Rh-3	DNA	Kenya2016	GF3Rh	GF3Rh-3	Rhizosphere	4575	1974
GF3Rh-4	DNA	Kenya2016	GF3Rh	GF3Rh-4	Rhizosphere	4575	1974
GF3Rh-5	DNA	Kenya2016	GF3Rh	GF3Rh-5	Rhizosphere	4575	1974
GF3S-1	DNA	Kenya2016	GF3S	GF3S-1	Soil	4575	1974
GF3S-2	DNA	Kenya2016	GF3S	GF3S-2	Soil	4575	1974
GF3S-3	DNA	Kenya2016	GF3S	GF3S-3	Soil	4575	1974
GF3S-4	DNA	Kenya2016	GF3S	GF3S-4	Soil	4575	1974
GF3S-5	DNA	Kenya2016	GF3S	GF3S-5	Soil	4575	1974
LG1-1	DNA	Kenya2016	S1LG1	LG1-1	Cryoconite	4875	2016
LG1-2	DNA	Kenya2016	S1LG1	LG1-2	Cryoconite	4875	2016
LG1-3	DNA	Kenya2016	S1LG1	LG1-3	Cryoconite	4875	2016
LG1-4	DNA	Kenya2016	S1LG1	LG1-4	Cryoconite	4875	2016
LG1-5	DNA	Kenya2016	S1LG1	LG1-5	Cryoconite	4875	2016
LG2-10-1	DNA	Kenya2016	S10LG2	LG2-10-1	Cryoconite	4750	2016
LG2-10-2	DNA	Kenya2016	S10LG2	LG2-10-2	Cryoconite	4750	2016
LG2-10-3	DNA	Kenya2016	S10LG2	LG2-10-3	Cryoconite	4750	2016
LG2-10-4	DNA	Kenya2016	S10LG2	LG2-10-4	Cryoconite	4750	2016
LG2-10-5	DNA	Kenya2016	S10LG2	LG2-10-5	Cryoconite	4750	2016
LG2-8-1	DNA	Kenya2016	S8LG2	LG2-8-1	Cryoconite	4675	2016
LG2-8-2	DNA	Kenya2016	S8LG2	LG2-8-2	Cryoconite	4675	2016
LG2-8-3	DNA	Kenya2016	S8LG2	LG2-8-3	Cryoconite	4675	2016
LG2-8-4	DNA	Kenya2016	S8LG2	LG2-8-4	Cryoconite	4675	2016
LG2-8-5	DNA	Kenya2016	S8LG2	LG2-8-5	Cryoconite	4675	2016
LG2-9-1	DNA	Kenya2016	S9LG2	LG2-9-1	Cryoconite	4700	2016
LG2-9-2	DNA	Kenya2016	S9LG2	LG2-9-2	Cryoconite	4700	2016
LG2-9-3	DNA	Kenya2016	S9LG2	LG2-9-3	Cryoconite	4700	2016
LG2-9-4	DNA	Kenya2016	S9LG2	LG2-9-4	Cryoconite	4700	2016
LG2-9-5	DNA	Kenya2016	S9LG2	LG2-9-5	Cryoconite	4700	2016
LGS2-1	DNA	Kenya2016	S2LG1	LGS2-1	Cryoconite	4875	2016
LGS2-2	DNA	Kenya2016	S2LG1	LGS2-2	Cryoconite	4875	2016
LGS2-3	DNA	Kenya2016	S2LG1	LGS2-3	Cryoconite	4875	2016
LGS2-4	DNA	Kenya2016	S2LG1	LGS2-4	Cryoconite	4875	2016
LGS2-5	DNA	Kenya2016	S2LG1	LGS2-5	Cryoconite	4875	2016
LGS4-1	DNA	Kenya2016	S4LG1	LGS4-1	Cryoconite	4875	2016
LGS4-2	DNA	Kenya2016	S4LG1	LGS4-2	Cryoconite	4875	2016
LGS4-3	DNA	Kenya2016	S4LG1	LGS4-3	Cryoconite	4875	2016
LGS4-4	DNA	Kenya2016	S4LG1	LGS4-4	Cryoconite	4875	2016
LGS4-5	DNA	Kenya2016	S4LG1	LGS4-5	Cryoconite	4875	2016
LGS6-1	DNA	Kenya2016	S6LG1	LGS6-1	Cryoconite	4875	2016
LGS6-2	DNA	Kenya2016	S6LG1	LGS6-2	Cryoconite	4875	2016
LGS6-3	DNA	Kenya2016	S6LG1	LGS6-3	Cryoconite	4875	2016
LGS6-4	DNA	Kenya2016	S6LG1	LGS6-4	Cryoconite	4875	2016
LGS6-5	DNA	Kenya2016	S6LG1	LGS6-5	Cryoconite	4875	2016
SF1-2	DNA	Kenya2016	SF1	SF1-2	Soil	4800	2004
SF1-3	DNA	Kenya2016	SF1	SF1-3	Soil	4800	2004
SF1-4	DNA	Kenya2016	SF1	SF1-4	Soil	4800	2004
SF1-5	DNA	Kenya2016	SF1	SF1-5	Soil	4800	2004
SF2Rh-1	DNA	Kenya2016	SF2R	SF2R-1	Rhizosphere	4800	1993
SF2Rh-2	DNA	Kenya2016	SF2R	SF2R-2	Rhizosphere	4800	1993
SF2Rh-3	DNA	Kenya2016	SF2R	SF2R-3	Rhizosphere	4800	1993
SF2Rh-4	DNA	Kenya2016	SF2R	SF2R-4	Rhizosphere	4800	1993
SF2Rh-5	DNA	Kenya2016	SF2R	SF2R-5	Rhizosphere	4800	1993
SF2S-1	DNA	Kenya2016	SF2S	SF2S-1	Soil	4800	1993
SF2S-2	DNA	Kenya2016	SF2S	SF2S-2	Soil	4800	1993
SF2S-3	DNA	Kenya2016	SF2S	SF2S-3	Soil	4800	1993
SF2S-4	DNA	Kenya2016	SF2S	SF2S-4	Soil	4800	1993
SF2S-5	DNA	Kenya2016	SF2S	SF2S-5	Soil	4800	1993
SF3Rh-1	DNA	Kenya2016	SF3R	SF3Rh-1	Rhizosphere	4800	1974
SF3Rh-2	DNA	Kenya2016	SF3R	SF3Rh-2	Rhizosphere	4800	1974
SF3Rh-3	DNA	Kenya2016	SF3R	SF3Rh-3	Rhizosphere	4800	1974
SF3Rh-4	DNA	Kenya2016	SF3R	SF3Rh-4	Rhizosphere	4800	1974
SF3Rh-5	DNA	Kenya2016	SF3R	SF3Rh-5	Rhizosphere	4800	1974
SF3S-1	DNA	Kenya2016	SF3S-2	SF3S-1	Soil	4800	1974
SF3S-2	DNA	Kenya2016	SF3S-2	SF3S-2	Soil	4800	1974
SF3S-3	DNA	Kenya2016	SF3S-2	SF3S-3	Soil	4800	1974
SF3S-4	DNA	Kenya2016	SF3S-2	SF3S-4	Soil	4800	1974
SF3S-5	DNA	Kenya2016	SF3S-2	SF3S-5	Soil	4800	1974

Appendix III: Gel imaging for the PCR products before the downstream analyses of purification and quantifications.



Appendix IV: Upstream analyses of Illumina sequence output

```
library(dada2); packageVersion("dada2")
path <- "/Volumes/G-tech_RAID/Mi_seq/Kenya16/output180807/" # CHANGE ME to
the directory containing the fastq files after unzipping.
list.files(path)

fnFs <- sort(list.files(path, pattern="_R1.fastq", full.names = TRUE))
fnRs <- sort(list.files(path, pattern="_R2.fastq", full.names = TRUE))
sample.names <- sapply(strsplit(basename(fnFs), "_"), `[`, 1)
plotQualityProfile(fnFs[1:2])
plotQualityProfile(fnRs[1:2])

filtFs <- file.path(path, "filtered", paste0(sample.names, "_F_filt.fastq.gz"))
filtRs <- file.path(path, "filtered", paste0(sample.names, "_R_filt.fastq.gz"))

out <- filterAndTrim(fnFs, filtFs, fnRs, filtRs, truncLen=c(250,210),
                     maxN=0, maxEE=c(1,3), truncQ=2, rm.phix=TRUE,
                     compress=TRUE, multithread=TRUE)
head(out)
errF <- learnErrors(filtFs, multithread=TRUE)
errR <- learnErrors(filtRs, multithread=TRUE)
plotErrors(errF, nominalQ=TRUE)
derepFs <- derepFastq(filtFs, verbose=TRUE)
derepRs <- derepFastq(filtRs, verbose=TRUE)
names(derepFs) <- sample.names
names(derepRs) <- sample.names
dadaFs <- dada(derepFs, err=errF, multithread=TRUE)
dadaRs <- dada(derepRs, err=errR, multithread=TRUE)
dadaFs[[1]]
mergers <- mergePairs(dadaFs, derepFs, dadaRs, derepRs, verbose=TRUE)
head(mergers[[1]])

seqtab <- makeSequenceTable(mergers)
dim(seqtab)
table(nchar(getSequences(seqtab)))

seqtab.nochim <- removeBimeraDenovo(seqtab, method="consensus",
multithread=TRUE, verbose=TRUE)
dim(seqtab.nochim)
sum(seqtab.nochim)/sum(seqtab)

getN <- function(x) sum(getUniques(x))
```

```
track <- cbind(out, sapply(dadaFs, getN), sapply(dadaRs, getN), sapply(mergers, getN),
rowSums(seqtab.nochim))
getN) with getN(dadaFs)
colnames(track) <- c("input", "filtered", "denoisedF", "denoisedR", "merged",
"nonchim")
rownames(track) <- sample.names
head(track)

##Save data
saveRDS(seqtab.nochim, "/Volumes/G-tech_RAID/Mi_seq/Kenya16/16Kenya.rds")
```

Appendix V: Subsequent analyses to prepare library for downstream analyses.

```
library(dada2)
Kenya <- readRDS("Kenya_plus.rds")
Ref_glacier <- readRDS("Refference_glacier_sequences.rds")

Kenya_all_glacier <- mergeSequenceTables(Kenya, Ref_glacier)

dim(Kenya_all_glacier)
table(nchar(getSequences(Kenya_all_glacier)))
Kenya_all_glacier_filtered <- Kenya_all_glacier[,nchar(colnames(Kenya_all_glacier))
%in% seq(402,428)]
dim(Kenya_all_glacier_filtered)
table(nchar(getSequences(Kenya_all_glacier_filtered)))

#Negative remove: see more https://github.com/benjjneb/dada2/issues/114
rownames(Kenya_all_glacier_filtered)
nega.samples <- c("Negative1", "Negative2") # CHANGE to names of your negative
controls
found.nega.samples <- colSums(Kenya_all_glacier_filtered[nega.samples,])>0
Kenya_all_glacier_filtered <- Kenya_all_glacier_filtered[,!found.nega.samples]
rownames(Kenya_all_glacier_filtered)

#Check negatives were removed correctly
rowSums(Kenya_all_glacier_filtered) # nega is zero
Kenya_all_glacier_filtered <- Kenya_all_glacier_filtered[-which
(rownames(Kenya_all_glacier_filtered) %in% c("Negative1", "Negative2")),]
rownames(Kenya_all_glacier_filtered)

#Create FASTA
asv_seqs <- colnames(Kenya_all_glacier_filtered)
asv_headers <- vector(dim(Kenya_all_glacier_filtered)[2], mode="character")
asv_fasta <- c(rbind(asv_headers, asv_seqs))
write(asv_fasta, "rep_set.fasta")

#Silva 132 data is available from
https://figshare.com/account/projects/60689/articles/7794296
taxa <- assignTaxonomy(Kenya_all_glacier_filtered, "silva_nr_v132_train_set.fa.gz",
multithread=TRUE)
write.table(taxa, "dada2_taxa.txt", sep='\t', row.names=FALSE, quote=FALSE)
##Seq Table
Kenya_all_glacier_filtered_t <- t(Kenya_all_glacier_filtered)#transpose the table
```

```
Kenya_all_glacier_filtered_t <- cbind('#OTUID' =
rownames(Kenya_all_glacier_filtered_t), Kenya_all_glacier_filtered_t)#Add '#OTUID'
to the header (required by biom)
write.table(Kenya_all_glacier_filtered_t, "dada2_seq_table.txt", sep='\t',
row.names=FALSE, quote=FALSE)
#Import meta data##
Meta <- read.delim("Metadata190304.txt", row.names = 1)
library(phyloseq)
library(ggplot2)
##Mke phyloseq data table
Glacier_all <- phyloseq(otu_table(Kenya_all_glacier_filtered, taxa_are_rows=FALSE),
                    sample_data(Meta),
                    tax_table(taxa))
Glacier_all
##remove Chloroplast and Mitochondria
Glacier_all_no_chloro<- subset_taxa(Glacier_all,
!Order=="Chloroplast"&!Family=="Mitochondria")
get_taxa_unique(Glacier_all_no_chloro, "Order")
get_taxa_unique(Glacier_all_no_chloro, "Family")

##Change row name from sequence to consequtive number
new.names <- paste0("ASV", seq(ntaxa(Glacier_all_no_chloro)))
seqs <- taxa_names(Glacier_all_no_chloro)
names(seqs) <- new.names
taxa_names(Glacier_all_no_chloro) <- new.names
#Pick up Kenyan glacier samples
###FOR KUJA## You can changes this setting as you want, here pick Kenya 16 from
glacier
Kenya2016_glacier = subset_samples(Glacier_all_no_chloro, Region=="Kenya2016" &
Type=="Glacier" & DNA=="DNA")
##general information
sample_names(Kenya2016_glacier)
otu_table(Kenya2016_glacier)[1:5, 1:5]
rank_names(Kenya2016_glacier)
tax_table(Kenya2016_glacier)
sample_variables(Kenya2016_glacier)
##Save the data
saveRDS(Kenya2016_glacier, "Kenya2016_glacier.rds")
```

Appendix VI: Taxonomic classifications using library Phyloseq

```
library(phyloseq)

install.packages("tidyverse")
library(tidyverse)

Kenya2016_glacier <- readRDS("Kenya2016_glacier.rds")

##Agglomerate at phylum level and transform to relative abundance
Kenya2016_glacier_phylum <- Kenya2016_glacier %>%
  tax_glom(taxrank = "Phylum") %>%
  transform_sample_counts(function(x) {x/sum(x)} ) %>%
  psmelt() %>%
  arrange(Phylum) %>%
  filter(Type=="Glacier" & DNA == "DNA") %>%
  filter(Abundance > 0.01) #if put rare taxa, bar chart become very messy (too much
taxonomy)

##Tax plot from all description sites
ggplot(Kenya2016_glacier_phylum, aes(x = Description, y = Abundance, fill =
Phylum)) +
  geom_bar(stat = "identity") +
  scale_fill_brewer(palette="Paired") +
  theme_bw() +
  guides(fill = guide_legend(reverse = TRUE, keywidth = 1, keyheight = 1)) +
  ylab("Relative Abundance (Phyla > 1%) ") +
  ggtitle("Kenyan glacier Phylum") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0.55, size=11),
        strip.text.x = element_text(size=11, face="bold", color ="black"),
        axis.title.x = element_blank())
scale_fill_manual(values =
c("MJ"="#a6cee3","ME"="#b2df8a","VE"="#33a02c","SUN"="#fb9a99","SCH"="#e31
a1c","NT"="#fdbf6f","SY"="#cab2d6","SER"="#ff7f00","BD"="#6a3d9a","HU"="#ffff
99","Precipitation"="#4eb3d3")) +
(https://stackoverflow.com/questions/28006281/ggplot-nested-x-axis-for-interaction-
factor-variables-in-bar-plot),
  #but I couldn't find good one to show (for our data, e.fg. indicate glacier and soil) or
put label manually by illustrator


  ##Tax plot from sites (average of 5 replicated samples)
```

```
  Kenya2016_glacier_phylum_mean <- aggregate(x=Kenya2016_glacier_phylum,
by=list(PHYLUM=Kenya2016_glacier_phylum$Phylum,
SITE=Kenya2016_glacier_phylum$Site), mean)

ggplot(Kenya2016_glacier_phylum_mean, aes(x = SITE, y = Abundance, fill =
PHYLUM)) +
  geom_bar(stat = "identity") +
  scale_fill_brewer(palette="Paired") +
  theme_bw() +
  guides(fill = guide_legend(reverse = TRUE, keywidth = 1, keyheight = 1)) +
  ylab("Relative Abundance (Phyla > 1%) \n") +
  ggtitle("Global glacier Phylum") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0.55, size=11),
        strip.text.x = element_text(size=11, face="bold", color ="black"),
        axis.title.x = element_blank())


##i.e Only Phylum_cyanobacteria
Kenya2016_glacier_cya <- subset_taxa(Kenya2016_glacier, Phylum=="Cyanobacteria")

#Agglomerate at genus level and transform to relative abundance
Kenya2016_glacier_cya_Genus <- Kenya2016_glacier_cya %>%
  tax_glom(taxrank = "Genus") %>%
  transform_sample_counts(function(x) {x/sum(x)} ) %>%
  psmelt() %>%
  arrange(Phylum) %>%
  filter(Type=="Glacier" & DNA == "DNA") %>%
  filter(Abundance > 0.01)

##Tax plot from all description sites
ggplot(Kenya2016_glacier_cya_Genus, aes(x = Description, y = Abundance, fill =
Genus)) +
  geom_bar(stat = "identity") +
  scale_fill_brewer(palette="Paired") +
  theme_bw() +
  guides(fill = guide_legend(reverse = TRUE, keywidth = 1, keyheight = 1)) +
  ylab("Relative Abundance (Phyla > 1%) ") +
  ggtitle("Kenyan glacier Phylum") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0.55, size=11),
        strip.text.x = element_text(size=11, face="bold", color ="black"),
        axis.title.x = element_blank())
```

```
##Tax plot from sites (average of 5 replicated samples)
Kenya2016_glacier_cya_Genus_mean <- aggregate(x=Kenya2016_glacier_cya_Genus,
by=list(GENUS=Kenya2016_glacier_cya_Genus$Genus,
SITE=Kenya2016_glacier_cya_Genus$Site), mean)

ggplot(Kenya2016_glacier_cya_Genus_mean, aes(x = SITE, y = Abundance, fill =
GENUS)) +
  geom_bar(stat = "identity") +
  scale_fill_brewer(palette="Paired") +
  theme_bw() +
  guides(fill = guide_legend(reverse = TRUE, keywidth = 1, keyheight = 1)) +
  ylab("Relative Abundance (Phyla > 1%) \n") +
  ggtitle("Global glacier Phylum") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0.55, size=11),
        strip.text.x = element_text(size=11, face="bold", color ="black"),
        axis.title.x = element_blank())
```

Appendix VII: Polygon ellipses for a relationship between the physicochemical parameters and the prominent genera distributions in the glacier.

```
Rcmdr> summary (ordiellipse(plot1, groups=site, conf=0.9,
kind='ehull'))
             LL         LL2          LL3           UL
UL2         UL3        UL4
NMDS1 0.1412913  0.1380548  0.39124186 -0.27686038 -0.4183
8310 -0.2195731 0.2889572
NMDS2 0.2261348 -0.0963429 -0.20848394 -0.10546228  0.0375
8635  0.1025951 0.2314181
Area  0.1387012  0.3724582  0.02745532  0.04735126  0.0457
9429  0.1500854 0.1696820

Rcmdr>  ordiareatest(plot1, groups=site, area='ellipse', k
ind='ehull', permutations=999)

Permutation test for the size of ordination ellipse
alternative hypothesis: observed area is smaller than rand
om ellipse

Permutation: free
Number of permutations: 999

      Area     5%    50% Pr(<sim)
LL  0.1387 0.1062 0.3817    0.087 .
LL2 0.3725 0.1116 0.3875    0.477
LL3 0.0275 0.1063 0.3943    0.005 **
UL  0.0474 0.0952 0.3764    0.014 *
UL2 0.0458 0.0959 0.3833    0.014 *
UL3 0.1501 0.0839 0.4006    0.129
UL4 0.1697 0.0948 0.3836    0.144
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
' 1
```

Appendix VIII: Script for heatmap analyses based on the Bray-Curtis dissimilarities between microbial compositions.

```
#Sample Filter

setwd("C:/Users/use/Desktop/L2_L6/l2")
community<- read.csv("l2.csv", header = TRUE, sep=",")
library(dplyr)
l2_LG2DNA_16c = subset(community[grepl("LG2", community$Sites),])
write.csv(l2_LG2DNA_16c, "l2_LG2DNA_16c.csv", row.names=FALSE)
l2_LG2DNA_16c1 <- read.csv("l2_LG2DNA_16c.csv", header = TRUE, sep=",",
row.names=1)
l2_LG2DNA_16cS = colSums(l2_LG2DNA_16c1)
write.csv(l2_LG2DNA_16cS, "l2_LG2DNA_16cS.csv")
l2_LG2DNA_16cM = colMeans(l2_LG2DNA_16c1)
write.csv(l2_LG2DNA_16cM, "l2_LG2DNA_16cM.csv")

#Heatmap
community<- read.csv("Kenya_L21.csv", header = TRUE, sep=",", row.names=1)
sample<- read.csv("E:\\Documentsaman\\Back up\\NMDS\\kuja_sample.csv", header =
TRUE, sep=",", row.names=1)

library(RColorBrewer)
library(Heatplus)
hmfamsite<- read.csv("Kenya_L21.csv", header = TRUE, sep=",", row.names=1)
data.dist <- vegdist(hmfamsite, method = "bray")
col.clus <- hclust(data.dist, "aver")
data.dist.g <- vegdist(t(hmfamsite), method = "bray")
row.clus <- hclust(data.dist.g, "aver")
plot(annHeatmap2(as.matrix(t(hmfamsite)), col = colorRampPalette(c("lightyellow",
"red"), space = "rgb")(850), breaks = 716, dendrogram = list(Row = list(dendro =
as.dendrogram(row.clus)), Col = list(dendro = as.dendrogram(col.clus))), cluster =
list(Col = list(cuth = 0.46, col = brewer.pal(10, "Set2")), Row = list(cuth = 0.37, col =
brewer.pal(10, "Set2"))), legend = 3, labels = list(Row = list(nrow = 3), Col = list(nrow
= 5.7))))
setwd("C:/Users/use/Desktop/L2_L6")
```

Appendix IX: Scripts for Rarefaction curves and Venn diagrams.

####Rarefaction curves

```
kujarare<- read.csv("otu.csv", row.names=1)
raremax <- min(rowSums(amanrare))
raremax
Srare <- rarefy(amanrare, raremax)
col <- c("blue", "red", "green", "brown", "violet", "black", "orange", "yellow",
"darkblue")
lty <- c("solid", "dashed")
lwd <- c(1, 2)
pars <- expand.grid(col = col, lty = lty, lwd = lwd, stringsAsFactors = FALSE)
out <- with(pars[1:26, ],
          rarecurve(amanrare, step = 20, sample = raremax, col = col,
                  lty = lty, label = TRUE))
Nmax <- sapply(out, function(x) max(attr(x, "Subsample")))
Smax <- sapply(out, max)
plot(c(1, max(Nmax)), c(1, max(Smax)), xlab = "Sample Size",
    ylab = "Species", type = "n")
abline(v = raremax)
for (i in seq_along(out)) {
  N <- attr(out[[i]], "Subsample")
  with(pars, lines(N, out[[i]], col = col[i], lty = lty[i], lwd = lwd[i]))
}
```

####Venn Diagrams

```
library(Vennerable)
ervenn<- read.csv("C:\\Users\\kuja\\Documents \\venn.csv", header = TRUE, sep=",")
ervenn1 = Venn(ervenn)
ervenn1
plot(ervenn1, doWeights=T, Type = "ellipses")
venn(ervenn, universe=NA, small=0.7, showSetLogicLabel=FALSE, simplify=FALSE,
show.plot=TRUE)
```

Appendix X: Scripts for Non-metric multidimensional scaling (NMDS) based on Bray-Curtis dissimilarities between microbial compositions of 16S rDNA.

```
```{r echo=FALSE}
# include this code chunk as-is to set options
knitr::opts_chunk$set(comment=NA, prompt=TRUE, out.width=750, fig.height=8,
fig.width=8)
library(Rcmdr)
library(car)
library(RcmdrMisc)
```
```{r echo=FALSE}
# include this code chunk as-is to enable 3D graphs
library(rgl)
knitr::knit_hooks$set(webgl = hook_webgl)
```
```{r}
setwd("/Users/josiahkuja/Documents/Analyses/NMDS")
```
```{r}
community<- read.csv("phyla_2016.csv", header = TRUE, sep=",", row.names=1)
```
```{r}
sample<- read.csv("sample_phyla_2016.csv", header = TRUE, sep=",",
  row.names=1)
```
```{r}
check.datasets(community, sample)
```
```{r}
dist.eval(community,'bray')
```
```{r}
Ordination.model1 <- metaMDS(community, distance='bray', k=2, trymax=999,
  autotransform=T, noshare=0.1, expand=T, trace=1, plot=F)
```
```{r}
check.ordiscores(community, Ordination.model1, check.species=F)
```
```{r}
Ordination.model1
```
```

```{r}
goodness(Ordination.model1)
```
```{r}
par(cex=1)
```
```{r}
plot1 <- plot(Ordination.model1, choices=c(1,2))
```
```{r}
abline(h = 0, lty = 3)
```
```{r}
abline(v = 0, lty = 3)
```
```{r}
attach(sample, warn.conflicts=F)
```
```{r}
ordisymbol(plot1, y=sample, factor='sampletype', legend=T,
  legend.x='topleft', legend.ncol=1, rainbow=T, cex=1)
```

Appendix XI: Scripts for Biplot, ANOSIM, ADONIS and Mantel tests using library Vegan.

```
###Biplot
community<- read.csv("E:\\Documentskuja\\NMDS\\water2.csv")
sample<- read.csv("E:\\Documentsaman\\NMDS\\sc_sample.csv", header = TRUE,
sep=",", row.names=1)

attach(community)
x = cbind(pH,EC,Albedo,Organic,Silicate,Carbonate)
summary(x)
cor(x)
pca1 = princomp(x, scores = TRUE, cor = TRUE)
biplot(pca1)

###Mantel
community<- read.csv("E:\\Back up\\NMDS\\genus.csv", header = TRUE, sep=",",
row.names=1)
sample<- read.csv("E:\\Back up\\NMDS\\sc_sample.csv", header = TRUE, sep=",",
row.names=1)
check.datasets(community, sample)
distmatrix1 <- vegdist(community,method='bray', na.rm=T)
dist.eval(community,'bray')
distmatrix2 <- vegdist(sample$salinity,method='euclidean')
mantel(distmatrix1, distmatrix2, method='pearson', permutations=999)

###ANOSIM
dune.dist <- vegdist(community)
attach(sample)
dune.ano <- anosim(dune.dist, SampleType)
summary(dune.ano)
plot(dune.ano)


###ADOSIN
data.adonis <- adonis(data.dist ~ pH + Slope + Altitude + Substrate, data=enviro)
data.adonis
Call:
  adonis(formula = data.dist ~ pH + Slope + Altitude + Substrate, data = enviro)

library(vegan)
```

Appendix XII: ANOSIM preliminary results and parameters

```
Rcmdr>  setwd("E:/Analyses/ANOSIM/Actinobacteria")

Call:
anosim(x = dune.dist, grouping = SampleType)
Dissimilarity: bray

ANOSIM statistic R: 0.7218
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0238 0.0344 0.0428 0.0598

Dissimilarity ranks between and within classes:
             0%     25%    50%     75%   100%    N
Between      89 1749.50 2496.0 3195.50 3901.5 2615
Cryoconite    1  155.50  344.0  667.00 3861.0  595
Rhizosphere 199  753.50 1109.5 1484.75 3103.0  300
Soil         18  808.75 1264.5 2056.00 3855.0  406


Call:
anosim(x = dune.dist, grouping = Age)
Dissimilarity: bray

ANOSIM statistic R: 0.6821
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0329 0.0474 0.0569 0.0712

Dissimilarity ranks between and within classes:
        0%     25%    50%     75%   100%    N
Between 89 1554.25 2381.5 3150.75 3901.5 2850
0        1  155.50  344.0  667.00 3861.0  595
12      73 1005.00 1579.0 1976.50 3750.0   91
23      98  792.00 1123.5 1563.25 3090.0  190
42      18 1029.50 1582.5 2306.25 3814.0  190
```

```
Call:
anosim(x = dune.dist, grouping = Altitude)
Dissimilarity: bray

ANOSIM statistic R: 0.4464
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0380 0.0539 0.0659 0.0829

Dissimilarity ranks between and within classes:
         0%     25%    50%     75%   100%    N
Between   1 1221.50 2150.0 3047.50 3901.5 3235
4575    249  790.00 1356.0 2265.00 3531.0   45
4600    307  789.00 1087.0 1537.00 2710.0   45
4675      6  988.00 1702.0 2797.00 3795.0  105
4700      8   80.25  275.5 2998.50 3562.0   10
4750     72  116.50  236.0  540.50  874.0   10
4800     18  764.50 1290.0 2289.75 3854.0  276
4875      4   89.75  233.5  475.75 1503.0  190


Call:
anosim(x = dune.dist, grouping = Site)
Dissimilarity: bray

ANOSIM statistic R: 0.5088
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0311 0.0436 0.0524 0.0743

Dissimilarity ranks between and within classes:
         0%     25%    50%     75%   100%    N
Between  49 1366.25 2333.5 3134.75 3901.5 2910
GF      249  988.50 1381.0 1864.00 3531.0  435
LL        1   72.00  255.0  699.00 3861.0  105
SF       18  764.50 1290.0 2289.75 3854.0  276
UL        4   89.75  233.5  475.75 1503.0  190
```

```
Call:
anosim(x = dune.dist, grouping = Site)
Dissimilarity: bray

ANOSIM statistic R: 0.5088
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0311 0.0436 0.0524 0.0743

Dissimilarity ranks between and within classes:
         0%     25%    50%     75%   100%    N
Between  49 1366.25 2333.5 3134.75 3901.5 2910
GF      249  988.50 1381.0 1864.00 3531.0  435
LL        1   72.00  255.0  699.00 3861.0  105
SF       18  764.50 1290.0 2289.75 3854.0  276
UL        4   89.75  233.5  475.75 1503.0  190


Call:
anosim(x = dune.dist, grouping = Site)
Dissimilarity: bray

ANOSIM statistic R: 0.5088
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0311 0.0436 0.0524 0.0743

Dissimilarity ranks between and within classes:
         0%     25%    50%     75%   100%    N
Between  49 1366.25 2333.5 3134.75 3901.5 2910
GF      249  988.50 1381.0 1864.00 3531.0  435
LL        1   72.00  255.0  699.00 3861.0  105
SF       18  764.50 1290.0 2289.75 3854.0  276
UL        4   89.75  233.5  475.75 1503.0  190
```

```
Rcmdr> community<- read.csv("level-6_Bacterio_assorted_ra.c
sv", header = TRUE, sep=",", row.names=1)

Rcmdr> sample<- read.csv("level-6_Bacterio_sample.csv", hea
der = TRUE, sep=",", row.names=1)

Rcmdr> community<- read.csv("level-6_Bacterio_assorted_ra.c
sv", header = TRUE, missing value = F, sep=",", row.names=1
)

Rcmdr> sample<- read.csv("level-6_Cyano_sample.csv", header
= TRUE, sep=",", row.names=1)

Rcmdr> community<- read.csv("level-6_Cyano_assort_ra.csv",
header = TRUE, sep=",", row.names=1)

Rcmdr> sample<- read.csv("level-6_Cyano_sample.csv", header
= TRUE, sep=",", row.names=1)

Call:
anosim(x = dune.dist, grouping = SampleType)
Dissimilarity: bray

ANOSIM statistic R: 0.4686
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0281 0.0389 0.0494 0.0657

Dissimilarity ranks between and within classes:
              0%     25%    50%    75%   100%    N
Between      1.0 1414.00 2089.0 3032.5 3032.5 2375
Cryoconite   2.0  205.00  441.0  683.5 1359.0  595
Rhizosphere 28.5 1192.50 2422.5 3032.5 3032.5  300
Soil        16.0 1100.75 1647.5 3032.5 3032.5  300


Call:
anosim(x = dune.dist, grouping = Age)
Dissimilarity: bray

ANOSIM statistic R: 0.5686
      Significance: 0.001
```

```
Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0424 0.0594 0.0742 0.0917

Dissimilarity ranks between and within classes:
          0%     25%    50%    75%   100%    N
Between    1 1393.25 2092.5 3032.5 3032.5 2574
0          2  205.00  441.0  683.5 1359.0  595
12        49  871.00 1569.0 3032.5 3032.5   91
23       106 1031.25 1863.5 3032.5 3032.5  190
42        16 1193.25 2020.0 2467.0 3032.5  120


Call:
anosim(x = dune.dist, grouping = Altitude)
Dissimilarity: bray

ANOSIM statistic R: 0.2374
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0416 0.0572 0.0688 0.0816

Dissimilarity ranks between and within classes:
            0%     25%    50%     75%   100%    N
Between    1.0 1034.00 1883.0 3032.50 3032.5 2975
4575      16.0  974.00 1530.0 1999.00 3032.5   45
4600     252.0  957.00 1234.0 3032.50 3032.5   45
4675      12.0  934.00 1622.0 3032.50 3032.5  105
4700      19.0  222.50  571.5  934.00 1146.0   10
4750       2.0   81.75  276.5  405.25  848.0   10
4800      28.5 1427.75 3032.5 3032.50 3032.5  190
4875       4.0  232.25  474.0  705.25 1133.0  190


Call:
anosim(x = dune.dist, grouping = Site)
Dissimilarity: bray

ANOSIM statistic R: 0.2562
      Significance: 0.001
```

```
Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0360 0.0462 0.0552 0.0729

Dissimilarity ranks between and within classes:
          0%     25%    50%     75%   100%    N
Between  1.0 1140.25 1947.5 3032.50 3032.5 2650
GF      16.0  963.00 1452.0 3032.50 3032.5  435
LL       2.0  116.00  367.0  597.00 1242.0  105
SF      28.5 1427.75 3032.5 3032.50 3032.5  190
UL       4.0  232.25  474.0  705.25 1133.0  190



Call:
anosim(x = dune.dist, grouping = Site)
Dissimilarity: bray

ANOSIM statistic R: 0.2562
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0360 0.0462 0.0552 0.0729

Dissimilarity ranks between and within classes:
          0%     25%    50%     75%   100%    N
Between  1.0 1140.25 1947.5 3032.50 3032.5 2650
GF      16.0  963.00 1452.0 3032.50 3032.5  435
LL       2.0  116.00  367.0  597.00 1242.0  105
SF      28.5 1427.75 3032.5 3032.50 3032.5  190
UL       4.0  232.25  474.0  705.25 1133.0  190


Call:
anosim(x = dune.dist, grouping = Site)
Dissimilarity: bray

ANOSIM statistic R: 0.2562
      Significance: 0.001
```

```
Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0360 0.0462 0.0552 0.0729

Dissimilarity ranks between and within classes:
          0%     25%    50%     75%   100%    N
Between  1.0 1140.25 1947.5 3032.50 3032.5 2650
GF      16.0  963.00 1452.0 3032.50 3032.5  435
LL       2.0  116.00  367.0  597.00 1242.0  105
SF      28.5 1427.75 3032.5 3032.50 3032.5  190
UL       4.0  232.25  474.0  705.25 1133.0  190



Rcmdr> dune.ano <- anosim(dune.dist, SampleType)

Rcmdr> summary(dune.ano)

Call:
anosim(x = dune.dist, grouping = SampleType)
Dissimilarity: bray

ANOSIM statistic R: 0.6827
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0260 0.0415 0.0525 0.0628

Dissimilarity ranks between and within classes:
            0%     25%    50%     75% 100%    N
Between     32 1633.50 2392.5 3051.75 3741 2502
Cryoconite   1  182.00  460.0 1261.00 3738  561
Rhizosphere 79  532.00  865.0 1370.50 2478  300
Soil        24  659.25 1125.0 1717.25 3540  378



Call:
anosim(x = dune.dist, grouping = Site)
Dissimilarity: bray
```

```
ANOSIM statistic R: 0.5853
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0274 0.0396 0.0513 0.0660

Dissimilarity ranks between and within classes:
        0%    25%  50%    75% 100%    N
Between 20 1376.0 2287 3018.0 3741 2777
GF      24  600.5  961 1418.5 2639  435
LL       5   92.0  324 2555.0 3733  105
SF      57  834.0 1336 1870.0 3650  253
UL       1  175.5  458 1009.0 2099  171



Rcmdr> dune.ano <- anosim(dune.dist, Age)

Rcmdr> summary(dune.ano)

Call:
anosim(x = dune.dist, grouping = Age)
Dissimilarity: bray

ANOSIM statistic R: 0.6307
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0347 0.0454 0.0536 0.0723

Dissimilarity ranks between and within classes:
        0%     25%    50%     75% 100%    N
Between 79 1445.25 2299.5 2998.75 3741 2722
0        1  182.00  460.0 1261.00 3738  561
12      24  791.25 1261.0 1798.00 3267   78
23      57  660.75 1037.5 1426.00 2591  190
42      32  629.25 1095.5 1608.50 2328  190


Rcmdr> plot(dune.ano)
```

```
Rcmdr> dune.ano <- anosim(dune.dist, Altitude)

Rcmdr> summary(dune.ano)

Call:
anosim(x = dune.dist, grouping = Altitude)
Dissimilarity: bray

ANOSIM statistic R: 0.4459
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0367 0.0503 0.0627 0.0778

Dissimilarity ranks between and within classes:
         0%     25%    50%     75% 100%    N
Between   9 1117.50 2104.5 2937.75 3741 3102
4575     32  401.00  758.0 1083.00 2008   45
4600    171  644.00 1022.0 1432.00 2533   45
4675      5  928.00 1851.0 2519.00 3740  105
4700      7   62.75  121.0 2999.75 3214   10
4750      8  145.50  461.5 1004.75 2036   10
4800     57  834.00 1336.0 1870.00 3650  253
4875      1  175.50  458.0 1009.00 2099  171



Rcmdr>  setwd("E:/Analyses/ANOSIM/Proteobacteria")

Call:
anosim(x = dune.dist, grouping = Age)
Dissimilarity: bray

ANOSIM statistic R: 0.681
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0341 0.0464 0.0631 0.0832
```

```
Dissimilarity ranks between and within classes:
         0%     25%    50%     75% 100%    N
Between 230 1266.00 1987.0 2677.00 3321 2457
0         1  104.25  206.5  355.75 2176  406
12      231  567.50  921.0 1475.50 3270   78
23       93  685.50 1138.0 1785.25 3222  190
42       44  711.50 1319.0 2109.50 3166  190

Call:
anosim(x = dune.dist, grouping = Altitude)
Dissimilarity: bray

ANOSIM statistic R: 0.389
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0396 0.0553 0.0720 0.0929

Dissimilarity ranks between and within classes:
         0%     25%    50%     75% 100%    N
Between   1 1008.50 1803.0 2577.50 3321 2751
4575    179  503.00 1124.0 1655.00 2506   45
4600    299  533.00  851.0 1490.00 2765   45
4675      5  496.00 1093.0 2265.50 3007   91
4700      2   30.25   52.0   68.50  132    6
4750     20   44.00  167.5  377.25  514   10
4800     44  671.00 1279.0 2376.00 3317  253
4875     10  129.50  217.5  350.00 1007  120

Call:
anosim(x = dune.dist, grouping = Site)
Dissimilarity: bray

ANOSIM statistic R: 0.469
      Significance: 0.001

Permutation: free
Number of permutations: 999

Upper quantiles of permutations (null model):
   90%    95%  97.5%    99%
0.0334 0.0474 0.0603 0.0741

Dissimilarity ranks between and within classes:
```

	0%	25%	50%	75%	100%	N
Between	45	1142.00	1942.0	2655.5	3321	2435
GF	179	650.50	1100.0	1709.5	3211	435
LL	1	32.25	81.0	233.5	1762	78
SF	44	671.00	1279.0	2376.0	3317	253
UL	10	129.50	217.5	350.0	1007	120

Appendix XIII: Difco™ R2A Agar Approximate Formula Per Litre

R2A Agar was developed by Reasoner and Geldreich1 for bacteriological plate counts of treated potable water. A low nutrient medium, such as R2A Agar, in combination with a lower incubation temperature and longer incubation time stimulates the growth of stressed and chlorine-tolerant bacteria. Nutritionally rich media, such as Plate Count Agar (Standard Methods Agar), support the growth of fast-growing bacteria but may suppress slow growing or stressed bacteria found in lower temperature conditions. When compared with nutritionally rich media, R2A Agar has been reported to improve the recovery of stressed and chlorine-tolerant bacteria from drinking water systems. R2A Agar is recommended in standard methods for pour plate, spread plate and membrane filter methods for heterotrophic plate counts.

Yeast Extract .....0.5 g

Proteose Peptone No. 3 .....0.5 g

Casamino Acids .....0.5 g

Dextrose .....0.5 g

Soluble Starch .....0.5 g

Sodium Pyruvate .....0.3 g

Dipotassium Phosphate .....0.3 g

Magnesium Sulfate .....0.05 g

Agar .....15.0 g

Appendix XIV: Isolation for the *Cyanobacteria* in BG-11 broth medium (units per litre)

Stock solutions for BG-11:	
Stock 1:	
Na_2MG EDTA	0.1g
Ferric ammonium citrate	0.6g
Citric acid. 1 H_2O	0.6g
$CaCl_2$. $2H_2O$	3.6g
Filter sterilize into a sterile bottle or autoclave	
Stock 2:	
$MgSO_4$. 7H2O	7.5g
Filter sterilize into a sterile bottle or autoclave	
Stock 3:	
K_2HPO_4 . 3 H_2O	4.0g
or K_2HPO_4	3.05g
Filter sterilize into a sterile bottle or autoclave	
Stock 5 (Microelements):	
H_3BO_3	2.86g
$MnCl_2$. 4 H_2O	1.81g
$ZnSO_4$. 7 H_2O	0.222g
$CuSO_4$. 5 H_2O	0.079g
$COCl_2$. 6 H_2O	0.050g
$NaMoO_4$. 2 H_2O	0.391g
or MoO_4 (85%)	0.018g

Appendix XV: Qubit analyses of the DNA quality

1	Conc.(ng/ul)	2	Conc.(ng/ul)	3	Conc.(ng/ul)	4	Conc.(ng/
S1LG1-1	58.2	S2LG1-2	46.4	S4LG1-3	50	S6LG1-4	56
S2LG1-1	47.4	S4LG1-2	41.8	S6LG1-3	60	S8LG2-4	57.6
S4LG1-1	57.4	S6LG1-2	60.6	S8LG2-3	70.2	S9LG2-4	64.8
S6LG1-1	52.4	S8LG2-2	46	S9LG2-3	59.8	S10LG2-4	50.4
S8LG2-1	50	S9LG2-2	54.8	S10LG2-3	68.2	S1LG1-5	43.6
S9LG2-1	53	S10LG2-2	51.6	S1LG1-4	62	S2LG1-5	57
S10LG2-1	60	S1LG1-3	48.6	S2LG1-4	55	S4LG1-5	32.2
S1LG1-2	75	S2LG1-3	67.8	S4LG1-4	70.4	S6LG1-5	50.4

6	Conc.(ng/ul)	7	8	Conc.(ng/ul)	9	Conc.(ng/ul)	10
GF2R-1	9.4	GF2S-2	GF2S-3	1.1	GF2S-4	26.2	GF2S-5
GF2S-1	25.8	SF2R-3	SF2R-4	35.4	SF2R-5	25.6	GF3R-1
SF2R-2	34	SF2S-3	SF2S-4	0.2	SF2S-5	2.5	GF3R-2
SF2S-2	1.5	SF3R-3	SF3R-4	35.6	SF3R-5	38	GF3R-3
SF3R-2	22	SF3S-3	SF3S-4	15.8	SF3S-5	15.5	GF3R-4
SF3S-2	11.4	GF1R-3	GF1R-4	35.8	GF1R-5	24.2	GF3R-5
GF1R-2	30	GF1S-3	GF1S-4	5.9	GF1S-5	31.6	GF3S-1
GF2R-2	13.7	GF2R-3	GF2R-4	52.8	GF2R-5	24.4	GF3S-2

Appendix XVI: Scripts for indicator species using function IndVal under package labdsv and vegan library

IndVal

```
library(labdsv)

samples <-read.delim("XXXX", row.names=1) #XXXX??????

iva <- indval(samples[,-1], samples[,1])

gr <- iva$maxcls[iva$pval<=0.01] #0.01?P-value???

iv <- iva$indcls[iva$pval<=0.01]

pv <- iva$pval[iva$pval<=0.01]

fr <- apply(samples[,-1]>0, 2, sum)[iva$pval<=0.01]

indvalsummary <- data.frame(group=gr, indval=iv, pvalue=pv, freq=fr)

indvalsummary <- indvalsummary[order(indvalsummary$group, -indvalsummary$indval),]

indvalsummary

write.table(indvalsummary, "output.txt", quote=F, col.names=F, append=T)
```

Appendix XVII: IndVal summary tables for the indicator species analyses output

Taxa	sites	indval	pvalue	freq
Verrucomicrobia; Luteolibacter	1	0.9557313	0.001	42
Actinobacteria; Mycobacterium	1	0.8950871	0.001	49
Gammaproteobacteria; Pseudomonas	1	0.8568845	0.001	35
Alphaproteobacteria; Amaricoccus	1	0.8445852	0.001	30
Verrucomicrobia; uncultured.10	1	0.8218437	0.001	48
Verrucomicrobia; Chthoniobacter	1	0.8007165	0.001	69
Alphaproteobacteria; Devosia	1	0.7995362	0.001	53
Actinobacteria; Nakamurella	1	0.7882489	0.001	86
Betaproteobacteriales; Variovorax	1	0.7839556	0.001	53
Bacteriodetes; Pedobacter	1	0.7837312	0.001	40
Bacteriodetes; Dyadobacter	1	0.7358664	0.001	20
Bacteriodetes; Chitinophagaceae; _	1	0.7086346	0.001	52
Alphaproteobacteria; Brevundimonas	1	0.7015072	0.001	30
Bacteriodetes; uncultured.2	1	0.6623283	0.001	56
Alphaproteobacteria; Sphingomonadaceae; _	1	0.6529637	0.001	59
Bacteriodetes; uncultured.3	1	0.6243061	0.001	44
Bacteriodetes; Flavobacterium	1	0.5981864	0.001	51
Actinobacteria; Friedmanniella	1	0.5960106	0.001	33
Bacteriodetes; uncultured.1	1	0.5944449	0.001	61
Deltaproteobacteria; uncultured	1	0.5886727	0.001	40
Alphaproteobacteria; Devosiaceae; _	1	0.5738906	0.001	35
Verrucomicrobia; Lacunisphaera	1	0.5372085	0.001	30
Alphaproteobacteria; Pedomicrobium	1	0.5176957	0.001	34
Verrucomicrobia; Opitutus	1	0.5126932	0.001	41
Actinobacteria; Micromonosporaceae; _	1	0.5111591	0.001	28
Actinobacteria; Conexibacter	1	0.5111058	0.001	61
Alphaproteobacteria; Rhizorhapis	1	0.5000848	0.001	59
Verrucomicrobia; Candidatus_Udaeobacter	2	0.7488299	0.001	41
Actinobacteria; uncultured.	2	0.7013924	0.001	42
Actinobacteria; Gaiella	2	0.6972303	0.001	52
Betaproteobacteriales; Noviherbaspirillum	2	0.6685557	0.001	44
Actinobacteria; uncultured	2	0.6653516	0.001	49
Alphaproteobacteria; Beijerinckiaceae; _	2	0.6391122	0.001	56
Alphaproteobacteria; Mesorhizobium	2	0.6260634	0.001	68
Actinobacteria; Solirubrobacterales; _	2	0.6154294	0.001	47
Actinobacteria; Solirubrobacter	2	0.5844472	0.001	38
Actinobacteria; Nocardioides	2	0.5712322	0.001	52
Actinobacteria; Modestobacter	2	0.5634162	0.001	39
Actinobacteria; Kineosporia	2	0.5595604	0.001	37
Bacteriodetes; Segetibacter	2	0.551632	0.002	46
Actinobacteria; Frankiales; _; _	2	0.5479532	0.001	53
Actinobacteria; Actinobacteria; _; _; _	2	0.5446309	0.001	61
Bacteriodetes; Flavisolibacter	2	0.5416717	0.001	36
Gemmatimonadetes; Gemmatimonas	2	0.537802	0.002	84
Gemmatimonadetes; uncultured	2	0.5342623	0.001	55
Alphaproteobacteria; Methylobacterium	2	0.5167974	0.001	42
Betaproteobacteriales; Methylotenera	2	0.5165552	0.001	59
Gemmatimonadetes; Gemmatimonadaceae; _	2	0.4980473	0.001	46
Betaproteobacteriales; _; _	2	0.4943025	0.001	36
Actinobacteria; Frankiales.bacterium	3	0.9132653	0.001	45
Cyanobacteria; uncultured	3	0.8727738	0.001	39
Alphaproteobacteria; Rhodovastum	3	0.8470015	0.001	55
Cyanobacteria; Tychonema	3	0.8398253	0.001	57
Alphaproteobacteria; Rhodobacter	3	0.7989983	0.001	43
Cyanobacteria; Oxyphotobacteria; _; _; _	3	0.7952862	0.001	50
Betaproteobacteriales; Rhizobacter	3	0.7849504	0.001	83
Bacteriodetes; Sporocytophaga	3	0.7817222	0.001	37
Bacteriodetes; Ferruginibacter	3	0.7673121	0.001	87
Actinobacteria; Oryzihumus	3	0.7626728	0.001	57
Bacteriodetes; Hymenobacter	3	0.7623592	0.001	73
Betaproteobacteriales; arcticum.group	3	0.7443208	0.001	46
Cyanobacteria; Phormidesmis	3	0.7391636	0.001	41
Alphaproteobacteria; Rhodovarius	3	0.7206288	0.001	37
Alphaproteobacteria; Polymorphobacter	3	0.7168662	0.001	60
Cyanobacteria; Chamaesiphon	3	0.6859159	0.001	48
Actinobacteria; uncultured.bacterium	3	0.65385	0.001	66
Actinobacteria; Angustibacter	3	0.6448577	0.001	62
Alphaproteobacteria; Acidisoma	3	0.6430815	0.001	36
Alphaproteobacteria; Acidiphilium	3	0.6264251	0.001	83
Actinobacteria; uncultured; _	3	0.6191876	0.001	68
Actinobacteria; uncultured..Actinobacteria.1	3	0.6041348	0.001	53
Actinobacteria; Rhodococcus	3	0.5852567	0.001	68
Alphaproteobacteria; Xanthobacteraceae	3	0.5479112	0.001	83
Bacteriodetes; Fibrella	3	0.5411786	0.001	54
Betaproteobacteriales; Polaromonas	3	0.5058109	0.001	80
Actinobacteria; Microtrichales; _.	3	0.4940394	0.002	55

Appendix XVIII: Scripts for the indicator species analyses based on the library ggplot and patchwork

```
p1 <- ggplot(genus72) + geom_point(aes(sites, Taxa))

p1

p1<- read.csv("~/Documents/Analyses/Indicator spp/output3.csv", stringsAsFactors = FALSE)

p1$sites <- as.character(p1$sites)

ggplot(p1, aes(x=sites, y=Taxa, size=indval, color=indval)) +
  geom_point(alpha=0.8) +
  scale_radius(range=c(1, 8)) +
  scale_colour_gradient(low = "#fff7f3", high = "#e31a1c") +
  xlab("Site") +
  labs( size = "IndVal??" ) +
  theme_classic() +
  theme( text = element_text(size=16),
        legend.justification = c("right", "bottom"), axis.text.y.right =element_blank())

p2 <- ggplot(genus72) +  geom_boxplot(aes(sites, Taxa, group = sites))

p2

p1 + p2
```