Distribution and characterization of bacterial communities in diverse Antarctic ecosystems by high-throughput sequencing

Bjorn Tytgat

Promotors

Prof. Dr. Anne Willems

Prof. Dr. Elie Verleyen

Dissertation submitted in fulfilment of the requirements for the degree of

Doctor (Ph.D) in Sciences, Biology (Ghent University)

Tytgat Bjorn | Distribution and characterization of bacterial communities in diverse Antarctic ecosystems by high-throughput sequencing

ISBN-number: 978-9-4619738-8-7

Funding: This thesis was supported by the Special Research Fund BOF of Ghent University (grant 0J28410) and by the Belgian Science Policy Office (BelSPO) project CCAMBIO.

Ph.D. Thesis, Faculty of Sciences, Ghent University, Ghent, Belgium

Publicly defended in Ghent, Belgium, April 28th, 2016

Examination committee

Prof. Dr. Koen Sabbe (Chairman)
Laboratory of Protistology and Aquatic Ecology
Department of Biology, Faculty of Sciences, Ghent University

Prof. Dr. Olivier De Clerck (Secretary)
Phycology Research Group, Ghent University
Department of Biology, Faculty of Sciences, Ghent University

Prof. Dr. Anne Willems (Promotor)
Laboratory of Microbiology
Department of Microbiology and Biochemistry, Faculty of Sciences, Ghent University

Prof. Dr. Elie Verleyen (Co-promotor)
Laboratory of Protistology and Aquatic Ecology
Department of Biology, Faculty of Sciences, Ghent University

Prof. Dr. Wim Vyverman
Laboratory of Protistology and Aquatic Ecology
Department of Biology, Faculty of Sciences, Ghent University

Prof. Dr. Aurélien Carlier
Laboratory of Microbiology
Department of Microbiology and Biochemistry, Faculty of Sciences, Ghent University

Prof. Dr. George A. Kowalchuk
Ecology & Biodiversity
Institute of Environmental Biology, Department of Biology, Utrecht University
The Netherlands

Dr. Arwyn Edwards
Centre for Environmental Microbiology
Microbiology Research Group
Institute of Biological, Rural and Environmental Sciences, Aberystwyth University
United Kingdom

Dankwoord

"*I am somehow less interested in the weight and convolutions of Einstein's brain than in the near certainty that people of equal talent have lived and died in cotton fields and sweatshops*."

- Stephen J. Gould, *The New Scientist, 1979*

"*We are like dwarfs on the shoulders of giants, so that we can see more than they, and things at a greater distance, not by virtue of any sharpness on sight on our part, or any physical distinction, but because we are carried high and raised up by their giant size*."

Bernard of Chartres, 8th century BP

"*I speak English well, I learn it from a book*"

Manuel & Guillaume

Niet dat ik claim om ook maar het talent van een Einstein of andere illustere onderzoekers te benaderen, verre van zelfs, maar ik ben zeer dankbaar om in een land en tijdperk te leven waarin ik – hoewel toch niet altijd zó evident – de kansen kon krijgen om de 'fouten' die ik als naïeve tiener gemaakt had recht te zetten, of om te studeren *tout court*, onafhankelijk van je afkomst of middelen. Dat ik hieraan zelfs een vervolg kon breien, met de bijkomende uitdagingen die het des te interessanter maakten, was een wens die uiteindelijk in vervulling is gegaan. Dit was echter niet mogelijk zonder de input, hulp, humor en babbels van en met verscheidene (over het algemeen ^^) betere specimens binnen onze soort.

Eerst en vooral moet – *en wil* – ik Prof. Dr. Anne Willems bedanken, zonder wie dit alles in eerste instantie niet mogelijk was. De vacature kwam op een cruciaal moment, en de bevestiging maakte er een bijzondere verjaardag van. Bedankt voor het gestelde vertrouwen voor dit, toch wel (een beetje), pionierswerk.

Je staat nu ook met naam en ~~toenaam~~ *titel* vereeuwigd in dit vodje papier, *Professor* Dr. Elie Verleyen. Bedankt voor de vele hulp, suggesties, discussies, avontuur Japan, en niet in het minst, het vertrouwen! Zonder jouw goedkeuren per Annes telefoontje kon ik hier niet aan beginnen. Je hebt je bureau weer (bijna) voor jou alleen ...

Prof. Dr. Wim Vyverman voor de discussies en adviezen, en samen met mijn promotoren, het vertrouwen. We gaan dan eens *jammen*. ;)

I wish to thank the members of the examination committee for their thorough review of the draft manuscript and helpful suggestions.

Het Nulde, chronologisch. Joke, mijn gids bij het betreden van de grote boze wereld, voor de babbels en de humor. Helen, voor de babbels, de discussies en de adviezen bij het eerste praktische werk. Pia, die het jaren met het prachtige uitzicht op mijn rug heeft moeten stellen. Bedankt voor de babbels en de hulp bij dat beangstigende prille labowerk (ARISA). Guillaume, partner in crime op de Ledeganck. De babbels, de discussies, de imitaties, de onbaatzuchtigheid. Een plezierige ramp om naast te zitten op conferenties ;). Anne Becker, half-en-half van het Nulde. Bedankt voor de fijne babbels en de morele steun de laatste maanden.

Het Sterrevolk: Maxime, partner in crime op de S8, voor de leute en de discussies, de scripts en de figuren die de thesis toch niet gehaald hebben. (Ook een ramp om naast te zitten tijdens conferenties ;)); Dagmar, om me te helpen bij de eerste stapjes van het NGS-verhaal; Sofie, voor de babbels (en *errug* interessante weetjes ;)) en technische/logistieke ondersteuning; Tine, om in te springen voor het verwerken van de RNA-extracties.

De technische back-up van het labo microbiologie: Margo, Liesbeth, Evie, Cindy, Epifania en Jeannine om me uit de nood te helpen (bij administratieve missers).

Although the meta-omics studies are not included in this thesis, I would also like to thank Prof. Satoshi Imura, 伊村先生, for the generous hospitality and provisioning of the facilities for the DNA and RNA-extractions at the National Institute for Polar Research, Tachikawa, Japan, and not in the least for demonstrating his skills with the ninja chopsticks. Also Dr. Takeshi Segawa for the aid and tips with the RNA-extractions, and occasional interpreting; Watanabe-san, for his effortless helping, despite the language barrier; Dr. Megumu Tsujimoto for collecting the samples according to our much demanding protocol; Dr. Akinori Takahashi for lending me his birding books; and the kind administrative assistants for the nice food and tea ceremony!

Het bestuur van Natuurpunt Haaltert, de mensen van plantenwerkgroep Allium en vogelwerkgroep Cinerea, die me de laatste twee jaar veelal hebben moeten missen tijdens activiteiten en vergaderingen, en die ik ondersteuningsgewijs nogal verwaarloosd heb. Morgen stop ik met uit te stellen!

Wim en Micheline, voor de vele hulp en opvang!

Opnieuw wil ik mijn moeder en jongste grote broer, Jurgen, bedanken voor de financiële en materiële steun tijdens mijn studies. Ook Dorine en Kurt voor de hulp de laatste jaren. Pa, ik hoop dat je trots zou geweest zijn. Ik mis je.

En zoals altijd, last but not least: Wendy, Zij-Die-Alles-Regelt, en met wie en door wie het allemaal begon. Zonder jou was ik waarschijnlijk nooit aan het TKO begonnen, laat staan de hogere studies. Mijn steun en toeverlaat, mijn beste vriendin. Dank je voor je geloof in mij en je liefde (en toch, ge zijt zot ...). Het laatste jaar was erg zwaar, maar we zijn er geraakt!
En ... Linnéa, door wie ik heel wat slaap (en waarschijnlijk deadlines ...) heb moeten missen, maar die het zonnetje in huis brengt. Je wordt een pracht van een vrouw, maar toch: you owe us! Wacht maar tot jij moet blokken ...

En nu, op naar nieuwe horizonten ... (en een gazonnetje!)

Bjorn

Contents

List of non-standard abbreviations and acronyms

a.s.l.	Above sea level
ACBR	Antarctic Conservation Biogeographic Region
ACC	Antarctic Circumpolar Current
AFL	ARISA Fragment Length
AR	Antarctic Region
ARISA	Automated Ribosomal Intergenic Spacer Analysis
bp	base pairs
BP	Before present
CCAMBIO	Climate Change and Antarctic Microbial BIOdiversity
CC	Climate Change
CFC	Chlorofluorocarbons
D_{BC}	Bray-Curtis dissimilarity
DGGE	Denaturing Gradient Gel Electrophoresis
DMS	Dimethylsulfide
EAIS	East Antarctic Ice Sheet
EC	Electric conductivity
GC	Global Change
GW	Global Warming
HGT	Horizontal gene transfer
HPLC	High Performance Liquid Chromatography
HTS	High-throughput sequencing
IC	Inorganic Carbon
Isl.	Island
kya	Thousand years ago
LGM	Last Glacial Maximum
mS/cm	milli-Siemens per centimeter
Ma	Mega-annum (million years)
MI	Marion Island
MDV	McMurdo Dry Valleys
Mya	Million years ago
MLSA	Multilocus sequence analysis
nt	nucleotides
NCBI	National Center for Biotechnology Information
NMDS	Non-metric Multidimensional Scaling

OTU	Operational Taxonomic Unit
PA	Presence/absence
PAR	Photosynthetically Active Radiation
PCR	Polymerase Chain Reaction
PFZ	Polar Frontal Zone
PP	Primary production
qPCR	quantitative Polymerase Chain Reaction
t-RFLP	Terminal Restriction Fragment Length Polymorphism
TAM	Transantarctic Mountains
TC	Total Carbon
TOC	Total Organic Carbon
VFH	Vestfold Hills
WAIS	West Antarctic Ice Sheet

Chapter 1. General introduction

1.1. The Antarctic Region

1.1.1 Antarctica: physical and environmental settings

Antarctica is Earth's southernmost continent. It is contained nearly completely within the Antarctic Polar Circle[1] (66°33'46.0") [2] (Bargagli, 2005), with only marginal areas situated in the Southern Temperate Zone directly north hereof, and hence, it contains the geographic South Pole (i.e., Earth's southern rotation axis' exit point). With about 14 million km² (Bargagli, 2005; Chown and Convey, 2007), it is the fifth largest continent (Cary et al., 2010), larger than Australia and Europe, representing about 10 % of the world's land surface (Kennedy, 1995) and 30 % of that in the Southern Hemisphere (Bargagli, 2005). Because of its high-latitude location and geographical isolation, the Antarctic continent is nearly completely covered by ice caps (Peck et al., 2005), containing 60 to 80 % of the Earth's freshwater stored in 90 % of the Earth's ice (Kennedy, 1993; Bargagli, 2005; Ugolini and Bockheim, 2008; Wilkins et al., 2013; Cavicchioli, 2015). In sharp contrast, only ~0.32 to 0.40 % of Antarctica's surface is estimated to be ice-free (Chown and Convey, 2007; Cary et al., 2010). The ice caps are on average 2160 m high (Cavicchioli, 2015), yet can locally exceed 4 km, reaching 4776 m in Adélie Land (Bargagli, 2005). Hence, Antarctica is considered to be the highest continent on Earth (Kennedy, 1995). It can globally be denominated the continent of extremes, or even, the continent of superlatives, since it is also the driest and windiest continent, has the largest seasonal variation in ice cover, with winter ocean ice reaching 19 x 10^6 km² (Wilkins et al., 2013), has the deepest continental shelf (Turner et al., 2009), and the largest wind-driven oceanic current (Bargagli, 2005; Wilkins et

[1] The Antarctic Circle is the northernmost latitude in the Southern Hemisphere at which the sun can remain continuously above or below the horizon for at least 24 hours

[2] As of December 7th, 2015 (http://www.neoprogrammics.com/obliquity_of_the_ecliptic/)

al., 2013). Despite receiving the highest amount of summer solar radiation of any continent, it is also the coldest place on Earth (Bargagli, 2005) and, has low annual levels of Photosynthetically Active Radiation (PAR; 400 – 700 nm) since it is covered in (partial) darkness from March until September (Laybourn-Parry, 2002; Laybourn-Parry and Wadham, 2014). Due to the presence of the ice sheets, the topography is characterized by a steep incline inland from the coastal zones, ending in the so-called Antarctic or Polar Plateau (ice sheet surface). Partly because of this difference in altitude, the mean annual temperatures are -50 °C on the plateau and -20 °C in coastal sites (Bargagli, 2005). The combination of low precipitation and low temperatures make the continent the largest cold desert on Earth (Chan et al., 2012).

1.1.2. Geography

The Antarctic continent can be divided into two main units: East and West Antarctica. These are physically separated by a 3500 km long mountain range, the Transantarctic Mountains (TAM) (Fig. 1.1). East Antarctica (or Greater or Continental Antarctica *sensu stricto*), covering an area of 10.35×10^6 km², is the largest unit, encompassing about two-thirds of the mainland (Steig and Orsi, 2013). It is covered by the East-Antarctic Ice Sheet (EAIS), which is on average 2300 m thick and has an estimated volume of nearly 29×10^6 km³ (Bargagli, 2005). Although its extent has fluctuated in the past, the EAIS persisted during the last 30 Ma (Huybers and Denton, 2008; Pollard and DeConto, 2009), partly because it nearly entirely rests on land mass (Christner et al., 2014). The TAM, stretching from Cape Adare at the Ross Sea (North Victoria Land) to the Pensacola Mountains (Queen Elisabeth Land), abut onto the Weddell Sea and are also considered to be part of East Antarctica. They are one of the longest mountain ranges on Earth, although several interruptions divide them into smaller mountain groups or ranges (Turner et al., 2009). Contrary to East Antarctica, West or Lesser

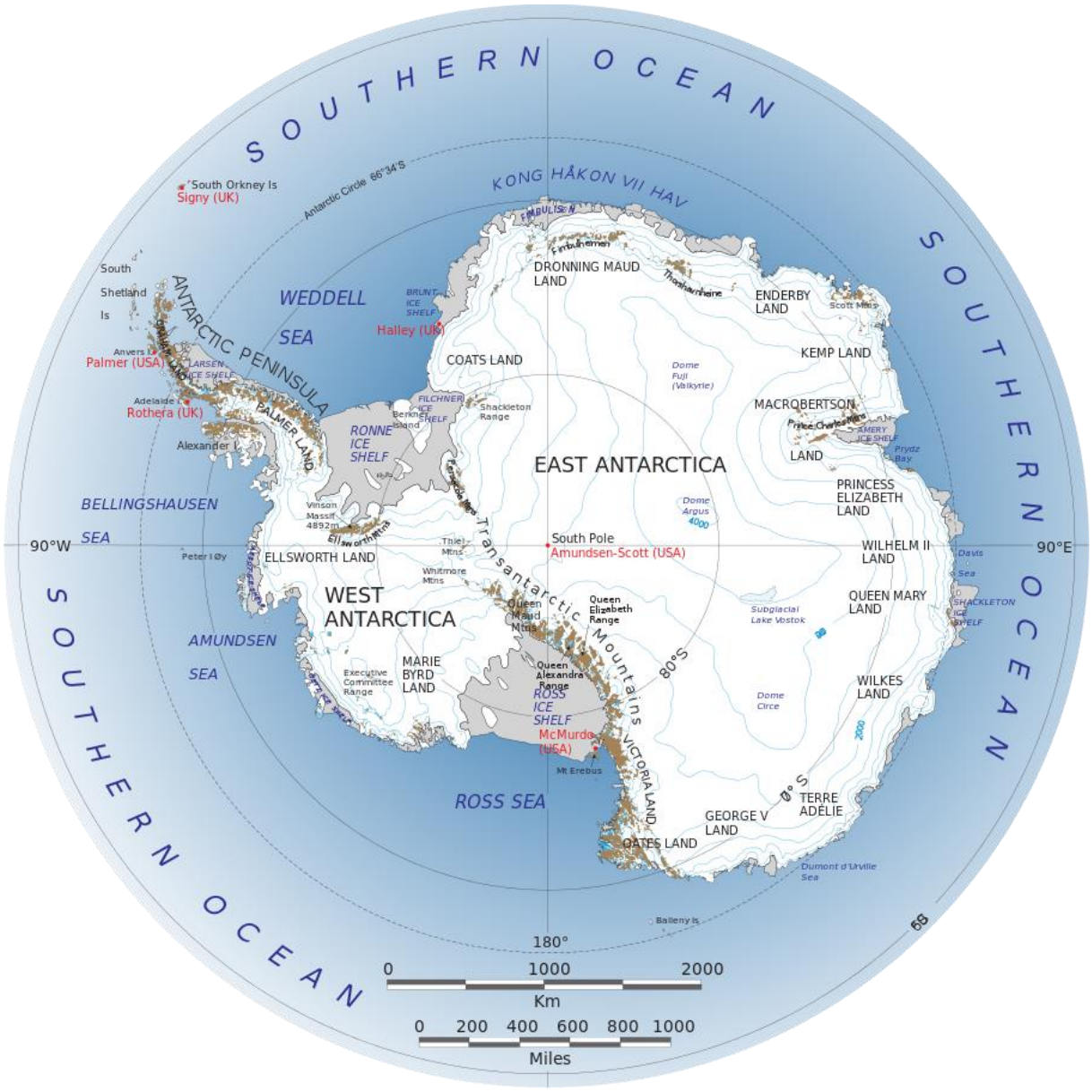

Figure 1.1. Overview of the Antarctic continent, showing the geographic subdivisions and several geographic landmarks (source: NASA). Ice-free areas are coloured brown.

Antarctica is mainly covered by an ice sheet (the West-Antarctic Ice Sheet, WAIS) that rests on bedrock which is situated below sea level. It is also far smaller than the EAIS (1.97×10^6 km²) and is on average only 850 m high (Turner et al., 2009). The WAIS has known considerable fluctuation in volume and extent for the last 10 Ma (Naish et al., 2009), and since it is a marine ice sheet, it is considered to be inherently unstable (Turner et al., 2009) and displays high melting rates (Steig and Orsi, 2013; Paolo et al., 2015). The mainland Antarctic ice sheet volume is estimated to be between 25.4 and 28.5 million km³ (Lythe et al.,

2001; Bargagli, 2005). Global water level equivalents of the WAIS and EAIS are estimated to be 5 m and 52 m, respectively (Lythe et al., 2001). The floating marine ice shelves surrounding these continental ice sheets limit the grounded ice sheet flow, slowing the continental ice discharge to the ocean (Paolo et al., 2015). The continental ice caps move seawards by on average 500 m per year, equalling 20 million km² per year, and calve off into the Southern Ocean (Bargagli, 2005). Observations from the last 18 years have shown an accelerated West Antarctic loss of the marine shelves up to 310 ± 74 km³ per year, or an increase of 70 % in the last decade. Some ice shelves in the Bellingshausen and Amundsen Seas have lost 18 % of their thickness in under two decades (Paolo et al., 2015). Ice shelves in East Antarctica have grown during 1994-2003, although now this volume gain apparently has ceased, and has since turned into a moderate loss (Paolo et al., 2015).

The Antarctic Peninsula is a bedrock archipelago (Bargagli, 2005), and is the only part of the continent that extends a significant way northwards from the main ice sheet, reaching 1200 km north into the Southern Ocean down to 63° S. These islands are separated by deep channels and joined by a grounded ice sheet, and cover an area of 0.52×10^6 km² (Turner et al., 2009). The Peninsula is separated from South America, about 1000 km further north, by the Drake Passage. It is considered to be part of Lesser Antarctica, although, just like the Sub-Antarctic and Continental Antarctica, it is often referred to as a separate biogeographic region (Turner et al., 2009; Chong et al., 2015). With an average width of 70 km and a mean height of 1500 m, the Peninsula acts as a barrier, affecting atmospheric circulation on the west-east trending oceanic and atmospheric circulations of the high southern latitudes, resulting in marked differences between its west and east coast (Bargagli, 2005; Turner et al., 2009). Because of the relatively warm and moist westerly winds, the western part of the Peninsula knows a milder maritime climate, with higher temperatures and more precipitation, which sporadically falls as rain during summer in the north (Bargagli, 2005). The western Peninsula

down to approximately 70° S, including the islands ranging from the South Sandwich Islands and Bouvetøya through the South Orkney and South Shetland Islands, is therefore referred to as the Maritime Antarctica semi-desert region (Ellis-Evans, 1996).

The Sub-Antarctic region is mainly oceanic, ranging from approximately 46° to 55—60° S, close to the Antarctic polar frontal zone (PFZ) (Chong et al., 2015). It comprises many islands or island groups, of which the more known ones are the Crozet Islands (Îles Crozet), the Kerguelen Islands, the Prince Edward Islands (including Marion Island), the South Georgia Group and Macquarie Island. Conditions are much milder, with higher average and less variable temperatures and precipitation rates compared to the Maritime and Continental Antarctic regions (Chong et al., 2015).

1.1.3. Climate and climate change

The geographical isolation of Antarctica and the Antarctic circumpolar current (ACC), its permanent ice cover, high average altitude, the long polar night and aphelion during the austral winter, and the low inclination of solar rays during the summer make it the coldest continent on Earth (Bargagli, 2005). The massive amount of ice present reflects the sun's radiation (albedo effect), and is an important buffer against global warming (Wilkins et al., 2013). In addition, the equator-to-polar difference in temperature results in a poleward transport of heat, effectively making Antarctica the main heat sink of the Southern Hemisphere (Bargagli, 2005).

Average summer temperatures on the continent are about -30 °C on the Plateau and -4 °C near the coast (Bargagli, 2005). Summer temperatures in coastal areas can locally exceed 0 °C, with observations of soil temperatures reaching *in extremis* 27.5 °C (Cowan and Ah Tow, 2004).

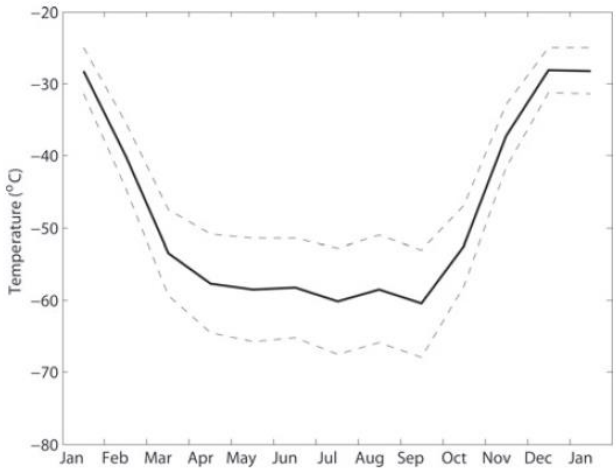

Figure 1.2. Monthly mean 2 m atmospheric temperature for the South Pole in the period 1994 to 2003. The dashed curves show the standard deviation of daily average temperatures about the monthly mean (taken from Town et al., 2008). The coreless winter can be noticed as the quasi stable temperature between March and September.

After a short summer period lasting only 6 to 8 weeks centred around early January, a rapid seasonal transition takes place, with a temperature drop of about 25 °C from the end of January until the end of March, resulting in a so-called "coreless winter" starting in April, which is characterized by minimal temperature variations for the next five months (Fig. 1.2). Average winter temperatures range from -25 to -70 °C, with the warmest temperatures in coastal locations, and the lowest on the Polar Plateau (Bargagli, 2005). The all-time lowest surface temperature recorded was -89.6 °C at the Vostok Station in July of 1983 (Convey, 1996; Bargagli, 2005). Typical annual values on the continent do not apply to the Antarctic Peninsula, which is at a lower latitude and over most of its length is divided into two distinct zones (Turner et al., 2009). The west coast of the Peninsula has a relatively mild maritime climate, and the annual average temperature (about -1.8 °C) is near 7 °C higher than that of the east coast at the same latitude. In fact, the latter zone is more affected by southerly winds and the greater extent of sea ice (i.e. the Ronne-Filchner Ice Shelf) (Bargagli, 2005). Temperatures in the Sub-Antarctic are much milder still, with average positive summer temperatures (Cowan and Ah Tow, 2004).

Because of the year-round low temperatures, all precipitation on the continent comes in the form of snow, yet, due to the extremely high aridity of the air, most precipitation sublimates

before reaching the surface, which is especially the case in the McMurdo Dry Valleys (MDV) (Cary et al., 2010). On average, 6 mm global sea level equivalent precipitates as snow on Antarctica each year, but the exact amount varies from region to region (Turner et al., 2009). The characteristically strong continental winds play an important role in snow redistribution, so that net accumulation in a particular place is also highly variable between years (e.g., Gorodetskaya et al., 2013, 2015).

The Maritime Antarctic, and to a lesser extent West Antarctica, is the fastest warming region on Earth since the 1950s (Turner et al., 2009; Bromwich et al., 2012; Steig and Orsi, 2013). The mean annual air temperature has risen by nearly 3 °C, and the mean winter air temperature by up to 7 °C. This has resulted in i.a. increased snowfall in West Antarctica and the western part of the Peninsula, negatively impacting Adélie penguin (*Pygoscelis adeliae*) breeding, a switch from snowfall to rainfall during summer to the north, enhanced primary production in Maritime Antarctic lakes, and the melting of glaciers (Quayle et al., 2002a; Turner et al., 2009). A decrease in sea ice has resulted in a decline of krill of up to 80 % in the last 30 years, additionally negatively impacting the West Antarctic Adélie penguins and other species (Trivelpiece et al., 2011; Ducklow et al., 2013).

Information on the EAIS is scarcer, but recent climate change effects are thought to be relatively modest in East Antarctica (Turner et al., 2009). The steep continental slope that runs between the ocean and the continental plateau prevents atmospheric eddies (the only significant source of heat to central Antarctica during the dark polar winter) from penetrating very far into the interior of East Antarctica (Steig and Orsi, 2013). There are however, distinct regional differences in for example snow accumulation rates (Monaghan et al., 2006). Local accretion of snow and ice has been observed, for example, in Law Dome (66°44’ S, 112°50’ E) near the coast (van Ommen and Morgan, 2010), while in the inner part of the

continent or in Dronning Maud Land, the snow volume has decreased between 1995 and 2004 (Monaghan et al., 2006).

1.1.3.1 Ozone hole and UV-radiation

Because of the poleward movement of the warmer air masses, Antarctica functions also as a sink for anthropogenic pollutants (Bargagli, 2005). A depletion of ozone levels in the stratosphere over Antarctica was noticed in the 1970s. Due to the use of chlorofluorocarbons (CFCs) and halons, nearly all ozone between a height of 14 and 22 km has disappeared, leaving the so-called ozone hole (Fig. 1.3). Ozone, as the only major atmospheric gas absorbing radiation at wavelengths of less than 300 nm, regulates the amount of ultraviolet radiation reaching the Earth's surface (Kennedy, 1995). Because of the ozone hole's polar location and the low presence of buffers such as clouds, 90 % of the radiation arriving at the top of the atmosphere over the ozone hole may reach the continental surface (Bargagli, 2005). Increased UV-B (280-315 nm) and UV-C (100-280 nm) radiation may have detrimental effects on microbiota, especially in terrestrial environments, which are less buffered compared to aquatic environments (Kennedy, 1995; Convey, 1996). These effects include mutagenesis, changes in membrane structure and interference with normal metabolic functioning, and may result in impaired growth, restricted mobility and changes in taxonomical composition in terrestrial communities (Kennedy, 1995). Also on flowering plants, increase of UV-radiation may have long term implications. For example, an increase in UV-radiation reduced the leaf size of *Deschampsia antarctica* Desv., while *Colobanthus quitensis* (Kunth) Bartl. showed an increased growth, and both produced more and larger seeds faster (Day et al., 1999).

UV-B irradiance towards the end of the 21st century is projected to be lower at mid to high latitudes by between 5 and 20 % respectively, and higher by 2 to 3 % in the low latitudes compared to 1980 (McKenzie et al., 2011). However, because of the ice-albedo feedback

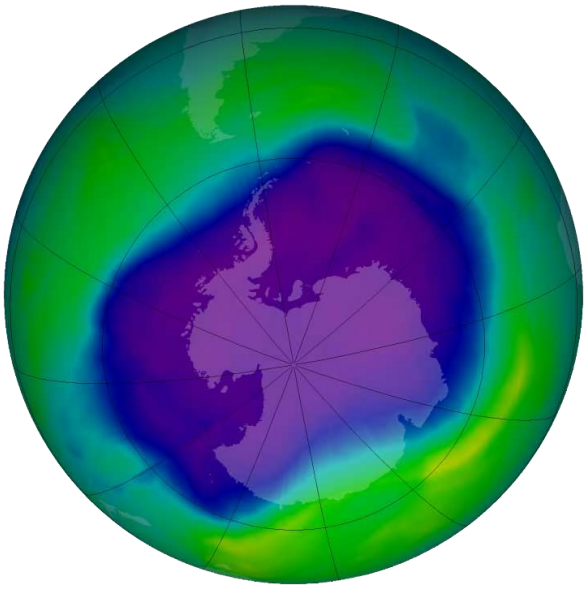

Figure 1.3. Image of the largest Antarctic ozone hole (purple) ever recorded (September 2006), over the Southern pole (source: NASA).

resulting from the melting of highly reflective ice surfaces and replacement by darker underlying surfaces, a greater warming at higher latitudes is predicted in response to increasing concentrations of greenhouse gasses (Bargagli, 2005). Ironically, the ozone hole allows reflected radiation to pass back to space and causes a heat loss, currently resulting in only a moderate warming of East Antarctica (Turner et al., 2009).

1.2. Geological and historical setting

Antarctica has known a long and dynamic geological history. Evidence of its origin as a relatively small cratonic nucleus centred on the Adélie Land regions of East Antarctica and the Gawler Craton region of South Australia dates back to the late Archaean (~2560 to 2510 Ma BP) (Boger, 2011). The protocontinent knew a quasi-continuous growth through accretions and collisions, forming a single large continent with mainly Proterozoic Australia and probably Laurentia (Boger, 2011). During several Ediacaran and Cambrian events, Laurentia rifted, creating the ancestral Pacific Ocean, while convergence with African and Indian landmasses resulted in the formation of the supercontinent Gondwana approximately 550 Mya. By ~250 Ma BP, nearly all landmasses on Earth, at that time represented by two

supercontinents, Gondwanaland (comprising South America, Africa, Antarctica, Madagascar, the Australian continent, New Caledonia, the Arab Peninsula and the Indian subcontinent) to the south, and Laurasia (comprising North America and Eurasia) to the north, had collided and had formed Pangaea. However, by ~213 Ma BP, Pangaea started to break up. This also initiated the disintegration of Gondwanaland itself, affecting sea currents, and changing the thermal regimes of both marine and terrestrial environments. By the Early Eocene (~50 Mya), after nearly 2500 Ma of shared history, Antarctica became separated from Australia through the opening of the Tasman Gateway, resulting in a near-complete deep water isolation of Antarctica and the establishment of the cool circumpolar Antarctic seaway (Boger, 2011). There is evidence from oxygen isotopes in marine sediments for a global cooling from 50 to 40 Ma BP, and several subsequent temperature drops resulted in the onset of a widespread Antarctic glaciation about 37─4 Ma BP (Lomolino et al., 2006; Bijl et al., 2013). Finally, the opening of the Drake Passage about 36-23 Mya resulted in the almost complete isolation of Antarctic marine biota. By the Middle Miocene (~15 Mya), the Antarctic Circumpolar Current was fully established, and the Polar Frontal Zone, a latitudinal temperature gradient, had developed, effectively separating the Southern Ocean from the other oceans.

Antarctica is likely to have shared much of its biota with South America and Australia, and possibly New Zealand (Peck et al., 2005; Waters and Craw, 2006), as is evident by the current disjunct distribution of e.g. southern beech (*Nothofagus* spp.) (Swenson et al., 2001) or marsupials (Metatheria) (Bargagli, 2005). Palaeontological evidence has indeed shown the presence of a *Nothofagus*-herb-moss tundra, and terrestrial (e.g., weevils and flies) and aquatic (e.g., the bivalve *Pisidium*, and at least one species of fish) macrofauna, persisting until possibly 12-2 Mya in the Transantarctic Mountains (Ashworth and Kuschel, 2003; Ashworth and Preece, 2003; Ashworth and Thompson, 2003; Stevens et al., 2006). While part of Gondwana, Antarctica was characterized by a temperate to sub-tropical climate, despite its

position at relatively high latitudes (Convey et al., 2008). Isolation from South America and subsequent cooling has led to the first formation of permanent ice-caps within the first few million years (Peck et al., 2005). However, during several warmer periods such as the Early Miocene (23-17 Mya), the continent hosted considerable expansion of cool temperate vegetation and fauna, most likely very similar to those in contemporary South America and New Zealand (Peck et al., 2005). After the warmer Early Pliocene (4.8-3.6 Mya), a sharp climatic cooling resulted in the expansion of the ice sheets, and the onset of the Pleistocene or Quaternary ice age. These fluctuations and intensification of the climatic extremes, led to episodic eradication of groups of organisms during the last 30 Ma (Peck et al., 2005), while the sea and wind currents probably largely prevented the influx of northerly propagules (e.g. Convey and Block, 1996).

1.3. Life on the frozen continent

Figure. 1.4. Wright Valley is one of the largest McMurdo Dry Valleys (South Victoria Land), a region long considered to be a sterile environment. (Photo courtesy of D. Saul)

While the general public often associates life on the Antarctic mainland with the apparently omnipresent and iconic penguins or seals and their relatives, these belong to the marine food chain and only come on land to breed or spend (considerable) time in the sun (Laybourn-Parry, 2002). Terrestrial life is less obvious, and parts of Antarctica, such as the McMurdo Dry

Valleys, were long considered to be adverse, sterile environments, or were thought to at most harbour very low levels of viable microbiota (Franzmann, 1996; Cowan and Ah Tow, 2004; Hogg et al., 2006; Cary et al., 2010). In recent years it has become clear that a macroflora dominated by lower plant groups (i.e. mosses, lichens and to a lesser extent, liverworts) can be found up to latitudes of 86° S (Peat et al., 2007). Only two species of flowering plants (*Deschampsia antarctica* E. Desv. (Poaceae) and *Colobanthus quitensis* (Kunth) Bartl. (Caryophyllaceae)) have been recorded, restricted to the western Antarctic Peninsula and the South Orkney and South Shetland Islands (Kennedy, 1995; Teixeira et al., 2010). Recent warming has increased the germination and establishment of seedlings (Smith, 1994) and these plants are expanding their range southwards (Parnikoza, 2011). Non-marine fauna is limited to a few invertebrate lineages, only two of which are insect species, both belonging to the Chironomidae (Diptera): *Belgica antarctica* or the Antarctic midge, an endemic wingless midge (Convey and Stevens, 2007) restricted to the Maritime region, and *Parochlus steinenii*, which has a wider distribution, including the high Andes and Tierra del Fuego, reaching its southern limits on the South Georgia and South Shetland Islands in Maritime Antarctica (Convey and Block, 1996). A third chironomid species (*Eretmoptera murphyi*) appears to have been accidentally introduced, suggesting at least for this group difficulty in colonizing the Antarctic (Convey and Block, 1996). Indeed, geographical and environmental isolation, and extremely selective conditions have produced a highly adapted and predominantly endemic biota (Peck et al., 2005; Chong et al., 2015). Contrary to Southern Ocean environments, the food webs of both terrestrial and non-marine aquatic habitats are very simple, with few trophic levels (Convey et al., 2008). Terrestrial faunal communities consist largely of micro-arthropods (mites, springtails), other microscopic invertebrates (nematodes, tardigrades, rotifers) and protozoans. Species diversity is low across all of these groups. As of 2014, 28 bdelloid rotifers, 66 monogonont rotifers, 59 tardigrades and 68 nematode species

have been recognized south of 60° S, while springtails and mites have a lower diversity (Velasco-Castrillón et al., 2014). Terrestrial food webs consist predominantly of assemblages of detritivores and microbivores, with very low contributions from true herbivores and predators. Micro-invertebrates include genera with characteristic trophic preferences (e.g., algivory, bacterivory, fungivory, nematophagy) (Convey et al., 2008). Continental faunal diversity is extremely low, with some of the simplest soil communities on Earth, in the most extreme environments containing three or less nematode species in two functional groups (Freckman and Virginia, 1997) or even no recorded fauna at all (Convey and McInnes, 2005).

Freshwater faunas are likewise extremely simplified, with a striking absence of fish (even on the Sub-Antarctic islands). Freshwater community structure in the maritime Antarctic includes single copepod and anostracan herbivores and a predatory copepod, in addition to benthic cladocerans and ostracods, and microscopic groups. Continental aquatic faunas are more reduced (Convey et al., 2008), with heterotrophic protists generally constituting the top level, while in some coastal lakes, a single crustacean species occurs as top predator (Wilkins et al., 2013).

Research has shown that life is, however, largely microscopic and these microorganisms are surprisingly diverse compared to the Antarctic plants and animals (Tindall, 2004), and can be active even in snow and ice (Carpenter et al., 2000; Lopatina et al., 2013). Moreover, quantitative studies have demonstrated that even McMurdo Dry Valley mineral soils harbour substantial populations of microorganisms with 10^6-10^8 cells g^{-1} of soil (Cowan et al. 2002) of high microbial phylogenetic diversity, dominated by *Bacteroidetes*, *Actinobacteria*, *Proteobacteria*, *Thermus-Deinococcus*, *Acidobacteria*, *Firmicutes* and *Cyanobacteria* (Aislabie et al., 2006; Smith et al., 2006). Microorganisms are fundamental to the functioning of Antarctic ecosystems, and while diversity is still relatively low compared to temperate

systems, metabolic flexibility is high, so that even few strains can provide the most necessary functions, such as carbon and nitrogen fixation (Wynn-Williams, 1996).

1.3.1. Adaptations to the Antarctic environment

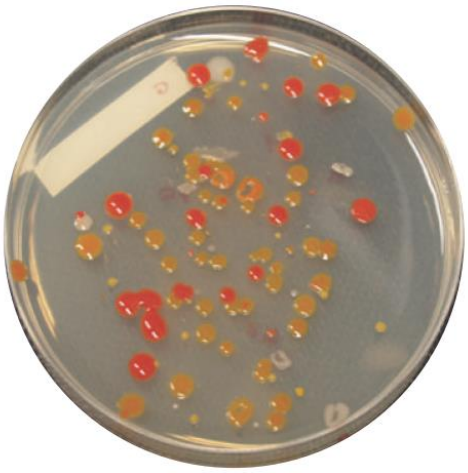

Figure 1.5. Heterotrophic bacteria isolates showing distinct pigmentation, providing protection against UV radiation (photo courtesy of Karolien Peeters).

The extreme conditions governing Antarctica have led to genetic, phenotypic and behavioural adaptations, enabling bacteria (and other organisms) to deal with the aridity, high UV-radiation and low temperatures. Terrestrial organisms, especially, have developed strategies to avoid or cope with these adverse conditions. They essentially include the colonization of lithic refuges (see section 1.4.1). Bacteria living in the less buffered upper mineral soils are subjected to the high UV-irradiance, and show distinct pigmentations to absorb the incident radiation, likely combined with highly developed DNA-repair systems, which has been demonstrated in e.g. *Deinococcus radiodurans* (Misra et al., 2013). Bacteria living in the subsurface permafrost layer generally lack pigmentation (Aislabie et al., 2006). Contrary to the upper layer bacteria, which are generally psychrotrophs (psychrotolerant) because of the need to cope with the thermal cycling of the soil, permafrost organisms are psychrophiles, thriving in a more stable environment at much lower temperatures (Margesin and Miteva, 2011). It has been shown that psychrophiles are able to reproduce at -12 °C, and are

metabolically active at temperatures of at least -20 °C (Margesin and Miteva, 2011). Functioning at low temperatures requires structural adaptation in proteins, conferring molecular flexibility, increasing catalytic efficiency and preventing cold-denaturation (D'Amico et al., 2006). Important adaptations are i.a. the increased production of (short) polyunsaturated fatty acids to maintain cell membrane fluidity, and enzymes which are able to function at low temperatures (Laybourn-Parry, 2002; Margesin and Miteva, 2011). However, the cold *in se* is not the limiting factor for life, rather the adverse effects of freezing water are. Psychrophiles produce various intra –or extracellular compounds to protect themselves against intracellular freezing or to minimize the deleterious effect of ice crystal formation (Margesin et al., 2005; Margesin and Miteva, 2011). For example, they produce antifreeze proteins, ice-binding proteins that have the ability to modify the ice crystal structure and inhibit the growth of ice. Aquatic organisms also produce high amounts of exopolymeric substances and antioxidant enzymes, since reactive oxygen radicals are well soluble in the cold water (Margesin and Miteva, 2011).

As a consequence, it has been hypothesized that the retrieval of organisms that are not adapted to these extreme conditions might indicate them to be opportunistic colonizers. For instance, obligate psychrophiles are more likely to be indigenous compared to psychrotolerants (Baker et al., 2003).

1.4. Antarctic ecosystems

Antarctica is a barren, desolate continent that because of its geographical isolation and relatively recent discovery - it was only first set foot on in 1821 - is often considered to harbour some of the last pristine ecosystems on Earth (Vincent, 2000; Baker et al., 2003), with only indirect anthropogenic influences such as the ozone hole (Laybourn-Parry and Wadham, 2014). Especially lakes are still considered to be highly pristine (Laybourn-Parry

and Wadham, 2014). Recently, however, increased research activities and tourism have resulted in intense direct effects on (microbial) communities through eutrophication, organic and inorganic contaminations (alien species, oils spills), and transportation of microbes between different sites, particularly near research stations, sampling sites and historical sites (Cowan, Chown, et al., 2011). However, still very little is known about the diversity and distribution of (indigenous) bacteria and their interactions with the highly variable Antarctic environment. There have been no large-scale, region-wide systematic surveys of the bacterial diversity, community structure and distribution. This is however urgently needed, given that (i) many regions and habitats (e.g. epiglacial lakes and cryconites) are still unexplored or under-sampled, (ii) the incidence of endemism appears to be high, (iii) some sectors of Antarctica are amongst the most rapidly warming regions on Earth (IPCC 2013), and (iv) habitats in these regions appear to respond quickly to climate changes (e.g., Quayle et al. 2002).

1.4.1. Terrestrial systems

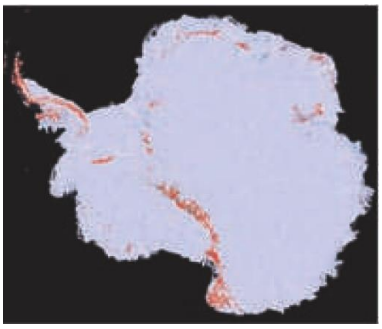

Figure 1.6. Map of Antarctica with contemporary ice-free areas highlighted in red (taken from Cary et al., 2010)

On the Antarctic continent, permanent ice-free regions – or oases – are scarce, and estimates of their cumulative surface vary from ‘less than 2 %’ (Cowan and Ah Tow, 2004) to the generally accepted figures of 0.32 % given by Chown and Convey (2007) or 0.40 % used by Cary et al. (2010) (Fig. 1.6). This constitutes respectively approximately 273,000 to as little as

44,000 km² of the total continental surface (Ugolini and Bockheim, 2008). Large areas of exposed soil are mainly situated along the edges of the continent: in low altitude coastal sites in the lower-latitude Antarctic Peninsula (7763 km²), parts of Eastern Antarctica (10,620 km²), including the Vestfold Hills, and in the largest contiguous range of ice-free regions (totalling 6861 km²), the McMurdo Dry Valleys (South Victoria Land); in mountain ranges such as the Ellsworth and Transantarctic Mountains (14,170 km²) and the North Victoria Land mountains (5412 km²); and additionally as scattered nunataks (1428 km²) (Cary et al., 2010).

The absence of vascular plants with their root systems (Ugolini and Bockheim, 2008), combined with centuries of wind and glacial dynamics, has depleted well-developed soil substrates, reducing the capacity to retain nutrients and water, while the absence of large numbers of primary producers has led to organic matter deprived, and thus oligotrophic, soils. The majority of the ice-free surface is hence mineral soil, composed of more or less fragmented local bedrock due to ages of freeze-thaw cycles, or debris carried from further inland through the seaward movement of glaciers (i.e. moraines). The general low water content in the mineral soils (< 2 %) further limits the buffering capacities, subjecting soil biota to i.a. desiccation, rapid temperature changes and high salt concentrations (Kennedy, 1993; Cary et al., 2010). Because of the low complexity food chains with apparently relatively few interactions, Antarctic terrestrial ecosystems are thought to be predominantly abiotically driven (Hogg et al., 2006). Nevertheless, terrestrial ecosystems, such as the MDV, are not homogeneous systems (Cowan and Ah Tow, 2004). Large physical and chemical gradients exist over small to broad geographical distances (Hogg et al., 2006; Cary et al., 2010), offering a wide range of potential microbial biotopes (Cowan and Ah Tow, 2004).

More important than temperature, liquid water is thought to be the prime limiting factor for biological activity (Kennedy, 1993; Convey and Smith, 2006). Water is provided in several ways. Direct water availability through precipitation is restricted, as water precipitates in a

Study site: the Sør Rondane Mountains

Figure 1.7. The eastern Sør Rondane Mountains (photo courtesy of René Robert)

In Chapters 2 and 3, the soil microbial diversity from the Sør Rondane Mountains (SRM) in eastern Dronning Maud Land, bordering Enderby Land, is reported. The SRM are a 220 km long east-west extending mountain range, situated about 200 km inland (Pattyn et al., 2010; Osanai et al., 2013), with peaks reaching 3300 m a.s.l. (Gorodetskaya et al., 2013). They consist of larger mountain subranges and more or less isolated nunataks. The Belgian Princess Elisabeth station was built on the ridge north of the Utsteinen nunatak, in the western part of SRM, several kilometres north of the main mountain range. The climatic conditions are rather mild, with average winter temperatures of -23 °C at Utsteinen, while summer averages are about -8 °C (Pattyn et al., 2010).

It is hypothesized that many terrestrial organisms or groups of organisms have persisted throughout at least the LGM in high altitude, inland ice-free refuges such as the SRM, and could possibly be remnants of an older Gondwanan biota (Stevens et al., 2006). While much research has been performed on the geology of the SRM (e.g. Matsuoka et al., 2006; Osanai et al., 2013), biological studies have only started since the construction of the Princess Elisabeth station in 2009. These studies have mainly focused on invertebrates (e.g. Stevens and D'Haese, 2014; Tsujimoto et al., 2014). Studies on bacterial diversity have been limited to a cultivation campaign of heterotrophic bacteria (Peeters, Ertz, et al., 2011) and Cyanobacteria (Namsaraev et al., 2010), while carbon and nitrogen fixation and light harvesting potential has recently been investigated through clone-libraries (Tahon et al., 2016a).

solid form, and either sublimates before reaching the surface or evaporates because of the dry katabatic winds rushing from the Plateau. During summer, meltwater runoff from glaciers wets adjacent soils, which may even saturate, and eventually drain into lakes or even rivers (Cowan and Ah Tow, 2004; McKnight et al., 2004). Additionally, a permafrost layer can typically be found within 30 cm of the soil surface, and liquid water is generally present between the permafrost and the surface, with most microbial activity and accumulation of organic matter occurring in this subsurface zone (Cary et al., 2010). Liquid water is potentially also available as a thin film on mineral surfaces at low temperatures (Dickinson and Rosen, 2003).

It is generally assumed that, besides liquid water, carbon and nitrogen are the limiting factors in terrestrial systems (Wada et al., 1981; Kennedy, 1993; Laybourn-Parry, 2002; Cowan, Sohm, et al., 2011). Multiple sources of organic matter and carbon are present, and include photoautotrophs such as mosses, lichens and macroscopic (e.g., *Prasiola crispa*) and microscopic algae and bacteria (e.g., diatoms and cyanobacteria). Their presence is usually scarce and patchy, and overall productivity levels are extremely low (for example, 1 to 20 g m^{-2} yr^{-1} in Taylor Valley, MDV) (Cary et al., 2010). Other carbon sources may have either a recent (e.g. lake inundations, sea spray and wind erosion of microbial mats) or a so-called legacy or relict organic matter origin (e.g., ancient glacial tills and lacustrine systems, marine incursions) which might represent the climatic history of a region (Wood, Rueckert, et al., 2008; Burkins et al., 2011). Carcasses of marine animals (e.g., seals; Fig. 1.8.) are very localized exogenous sources that support distinct microbial and invertebrate communities (Barrett, Virginia, Wall, et al., 2006). Nitrogen fixation is generally assumed to be performed by Cyanobacteria, in particular by members of the *Nostocales* order (Cowan, Sohm, et al., 2011; Tahon et al., 2016a), and N-fixing genes have also been shown to be more abundant in lichen-dominated plots (Yergeau, Kang, et al., 2007). Terrains which have been ice-free for

Figure 1.8. Mummified seals can be found scattered across the Dry Valleys and be hundreds of years old (photo courtesy of Mila Zinkova)

longer periods commonly have higher nitrogen concentrations due to a prolonged exposure to atmospheric nitrogen deposition (Cary et al., 2010; Magalhães et al., 2012). Nevertheless, the overall lack of precipitation precludes large-scale flushing of accumulated nitrogen and salts, leading to the unusual high salt content of Antarctic soils, which in extreme cases prevents liquid water to be biologically available for microorganisms (Kennedy, 1993). Such soils have indeed been shown to contain less diverse communities with simpler foodweb structures (Magalhães et al., 2012). In stark contrast, ornithogenic soils, enriched by mainly guano deposits in penguin colonies, are both nutrient-rich (i.e., high carbon, nitrogen and phosphorus content) and water-rich, and are generally high in microbial biomass (Aislabie et al., 2009). Nevertheless, such soils did not prove to be more diverse than mineral soils from the same region (Aislabie et al., 2009). Other nutrients (minerals, inorganic molecules) originate from aeolian or meltwater transport or physicochemical weathering of the local soil or bedrock (Barrett, Virginia, Hopkins, et al., 2006).

Additional environmental factors that impose important effects are the stability of soils and the exposure to wind and UV-irradiation (Cary et al., 2010). Mineral soils can furthermore experience multiple freeze-thaw cycles daily, with differences in temperatures of 20-50 °C or more in less than three hours (Convey, 1996; Cary et al., 2010), leading to a highly fragmented substratum (Convey, 1996) and potentially lethal conditions for soil microbiota

Figure 1.9. Cryptoendolith (photo courtesy of Guillaume Dargaud)

(Cowan and Ah Tow, 2004). To cope with these conditions, organisms have developed avoidance strategies, forming highly-specialized communities in lithic environments (Pointing et al., 2009a; Cowan et al., 2010). These habitats provide physical stability, thermal buffering and enhanced moisture availability over the surrounding soil, and protection against drought and wind, high UV-B radiation and excessive photosynthetically active radiation, and freeze-thaw-cycles (Chan et al., 2012). Hypolithic communities, or hypoliths, are microbial communities that essentially exist on the ventral side of rocks or pebbles, in particular translucent rocks, where light penetrates, providing energy for autotrophic organisms (Pointing et al., 2009a; Valverde et al., 2015). A survey in the MDV has shown that hypoliths are only found beneath quartz and marble rock with a surface of more than 4 cm^2 (Cowan et al., 2010). Hypolithic communities are thought to be the predominant primary producers in hyperarid environments (Cowan, Sohm, et al., 2011; Pointing and Belnap, 2012). They are generally considered to be dominated by *Cyanobacteria*, yet Cowan et al. (2010) classify these hypolithic communities into three categories: Type I (cyanobacteria dominated), Type II (fungi dominated) and Type III (moss dominated). These different types may reflect stages in community development (Cowan et al., 2010). Other lithic ecosystems are the relatively infrequent chasmoliths, which are microbial communities within cracks and fissures of rocks (Pointing 2009), while endoliths are consortia of microorganisms living inside the rocks themselves (Fig. 1.9). The latter are generally recovered from sandstone rocks, which have

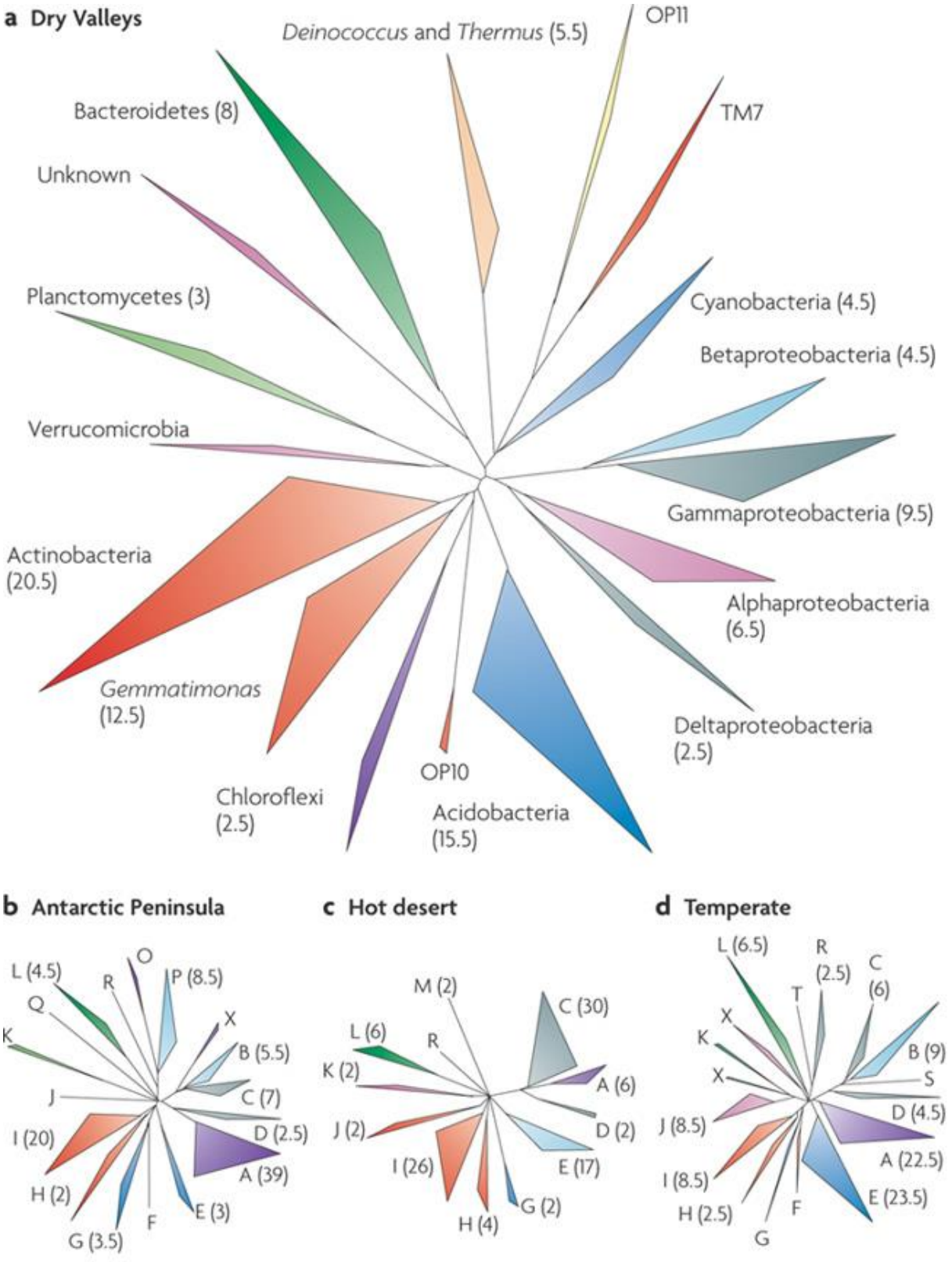

Figure 1.10. Phylogenetic diversity of bacterial 16S ribosomal RNA gene sequences from the McMurdo Dry Valleys (a), Antarctica, and comparison with bacterial diversity in clone libraries from the Antarctic Peninsula, a hot desert and temperate surface soils. (Taken from Cary et al. 2010)

relatively large pore sizes. In the Antarctic, endoliths are typically found up to 3 mm deep, where light is still able to penetrate, albeit highly attenuated (Walker and Pace, 2007).

The main bacterial phyla observed in Antarctic terrestrial systems – mainly from the MDV – (e.g., *Proteobacteria*, *Actinobacteria*, *Acidobacteria*, *Firmicutes*, *Bacteroidetes*, *Gemmatimonadetes*, *Chloroflexi* (Fig. 1.10)) are not very different to those from other parts of the world (Janssen, 2006). Clear exceptions are the *Cyanobacteria* and candidate division *FBP* (Lee et al., 2013). *Cyanobacteria* are arguably the most important primary producers in the oligotrophic Antarctic ecosystems, and some groups are known for nitrogen fixation as well. The candidate division *FBP* is highly intriguing, as representatives are apparently mainly recovered from cold associated areas.

No unambiguous conclusions can, however, be made about the correlation of environmental parameters with the occurrence of these phyla. For the generally described parameters (carbon or organic matter content, pH, conductivity, moisture content), both positive and negative correlations with community structure and composition have been reported (e.g., Niederberger et al., 2008; Magalhães et al., 2012; Stomeo et al., 2012; Geyer et al., 2014; Kim et al., 2015), with the clear exception of the positive correlation between moisture content and *Cyanobacteria* (e.g., Kim et al., 2015).

Inland terrestrial systems, such as the SRM, are, however, less well characterized, and it is not clear if the environmental gradients shaping microbial communities in such environments are similar as those in the MDV.

1.4.2. Aquatic ecosystems

Bioavailability of water is considered to be the most important limiting factor to life in Antarctica (Kennedy, 1993), effectively curtailing biological activity in terrestrial systems throughout most of the year (Laybourn-Parry, 2002). With Antarctica being the largest cold

desert on the planet, it is no wonder that most life on the continent is concentrated in those lakes where liquid water is present year-round (Laybourn-Parry, 2002). Indeed, the availability of liquid water stimulates the most productive and diverse microbial communities (Cowan and Ah Tow, 2004).

Antarctic aquatic ecosystems cover an array of different aquatic habitats, that each show a vast physical and chemical range. The marine environment is likely to be the better studied one (e.g., Peck et al., 2005; Ghiglione and Murray, 2012; Ghiglione et al., 2012; Carr et al., 2015). The Southern Ocean plays an important role in the Earth's climate and ocean currents, including the ACC, and is home to an incredible diversity of organisms (Wilkins et al., 2013), with many species of fish and invertebrates resulting from post-isolation radiation of few lineages (Peck et al., 2005). Hence, endemism is high, possibly even in bacteria (Landone Vescovo et al., 2014). Furthermore, microorganisms in the cold Southern Ocean play significant roles in global biogeochemical cycles, such as carbon sequestration (Cavicchioli, 2015).

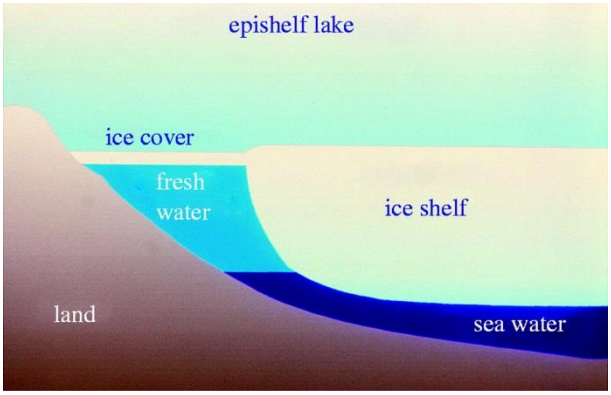

Figure 1.11. Epishelf lake (Taken from Laybourn-Parry & Pearce, 2007)

Epishelf lakes, which are freshwater columns floating on denser and colder seawater trapped between the ice shelf and land mass (Laybourn-Parry and Pearce, 2007), form an interface between marine and continental ecosystems (Fig. 1.11). Their basins continue below the ice

shelf and hence know a tidal regime. Despite being a dynamic system, no significant mixing of freshwater and seawater occurs. Epishelf lakes are rare (Mueller and Vincent, 2003; Laybourn-Parry and Pearce, 2007; Veillette et al., 2011), and are particularly vulnerable to climate change (Veillette et al., 2008). Relatively little is known about the biology of these ultra-oligotrophic lakes, but analysis of Beaver Lake showed very low bacterial and primary production (Laybourn-Parry, Quayle, et al., 2001).

Compared to the Southern Ocean, much less research has been performed on continental water bodies. On the Antarctic continent, numerous lakes and ponds can be found in the regions not covered by glacial ice (so-called oases), which are mainly situated in coastal lowland areas, such as the Vestfold Hills with an estimated 300 water bodies, the MDV, and the Maritime Antarctic (Laybourn-Parry and Wadham, 2014; Cavicchioli, 2015). Antarctic continental lakes show the highest diversity of lake types on the planet (Laybourn-Parry and Wadham, 2014). Continental waters vary from meltwater systems atop glaciers, which may lack nearly any ions or be sediment rich, up to hypersaline marine derived lakes, with salinities reaching ten times that of seawater (Laybourn-Parry, 2002; Laybourn-Parry and Wadham, 2014). Lakes can be permanently ice-free, such as hypersaline Deep Lake (Vestfold Hills, East Antarctica), to perennially covered, such as Lake Fryxell (MDV, South Victoria Land). Lakes can be holomictic, meromictic, monomictic or dimictic; aeriated or partly or completely anoxic; and show steep salinity, oxygen or chemical clines (Laybourn-Parry and Wadham, 2014). Because of many peculiarities, much attention has been given to high salinity lakes or lakes governed by other extreme conditions (Lauro et al., 2011; Yau et al., 2013; Laybourn-Parry and Bell, 2014).

Freshwater lakes or ponds can be formed through the collection of summer meltwater in catchment basins or in depressions abut onto glaciers (proglacial lakes) (Cowan and Ah Tow, 2004). Supraglacial systems are temporary or seasonally formed as either lenses of meltwater

or as cryoconite holes on top of glaciers or lake ice covers (Laybourn-Parry and Wadham, 2014). The latter originate from windblown dust deposits or larger rock material (from erosion) that absorb more solar energy than the ice, melting the surrounding ice and over time sinking into the ice. These holes form traps for other dust particles, increasing the absorbance and area of the cryoconite. The presence of liquid water and nutrients provided by the particulate deposits can sustain a microbial community. Cryoconites harbour diverse viral, archaeal, bacterial, microeukaryote and meiofaunal life (Anesio et al., 2009; Obbels et al., submitted), and remarkable rates of primary production and respiration, sometimes approaching those of temperate soils, have been associated with these ecosystems (Anesio et al., 2009).

The majority of coastal oasis lakes were formed by glacial processes and isostatic uplift (Laybourn-Parry and Pearce, 2007). Saline systems are either derived through marine incursions (e.g., Ace Lake), or because of the gradual increase of salt through the accumulation of sea spray or chemical weathering of rocks drained into a lake catchment (e.g., Lake Fryxell) (Cowan and Ah Tow, 2004). Sublimation and evaporation will then lead to hypersaline or brine lakes with salt concentrations up to ten times sea water (Laybourn-Parry and Wadham, 2014). This depresses the freezing-point, with extreme values down to -53 °C in the World's most saline natural water body, Don Juan Pond (MDV) (Dickson et al., 2013).

Most intriguing are probably subglacial lakes, which represent an enormous amount of unexplored territory. They impact the ice sheet movements and their sediments could provide invaluable information on past climate change (Pearce, 2009; Wright and Siegert, 2012). By the late 1960s, radio-echo sounding of Polar (Greenland) ice sheets was well established, and charting of Antarctic subglacial lakes started in 1966 (Robin et al., 1970; Oswald and Robin, 1973). Several inventories have since been performed (e.g., Siegert et al., 1996, 2005; Siegert, 2000), and the number of subglacial lakes is estimated to be 379 (Wright and Siegert, 2012).

They are among the most extreme viable systems (Wright and Siegert, 2012), with combined stresses of high pressure, permanent darkness, and low temperatures, nutrient availabilities and oxygen concentrations (Pearce, 2009). They likely harbour unique microbial assemblages, adapted to these extreme environments, including potentially unique relict and endemic organisms, as time since isolation or migration through the ice sheets can be thousands of years (Vincent, 2000; Pearce, 2009). However, although certain systems are indeed isolated hydrologically and can have mean water retaining times of several thousands of years, such as up to 13,000 years in Lake Vostok (A. Richter et al., 2014), other subglacial lakes can be connected, exchanging water, nutrients and most likely microorganisms (Wingham et al., 2006).

Like their terrestrial counterparts, lacustrine ecosystems are characterized by truncated foodwebs, although diversity is considerably higher and might have some additional trophic levels (Convey et al., 2008; Wilkins et al., 2013). Many marine derived saline lakes were formed from pockets of seawater trapped in closed basins when the land rose. The original eukaryotic marine communities became progressively simplified, leaving a few dominant nanoflagellates, dinoflagellates, ciliates and diatoms (Laybourn-Parry and Pearce, 2007). In these simple foodwebs, viruses appear to be important components enforcing a top-down control, promoting gene exchange and microbial evolution (Anesio and Bellas, 2011). Because of their vast numbers, they are pivotal in carbon cycling (Laybourn-Parry, Hofer, et al., 2001). Their importance has been supported by microscopy-based observations of viral density, virus-to-bacteria ratios, and infection rates that are higher in Antarctic lakes than lower-latitude systems (Laybourn-Parry, Hofer, et al., 2001; Säwström et al., 2007).

Antarctic lakes are very unproductive, with photosynthesis levels in the region of 0.5–30 µg l^{-1} day^{-1} because of the low annual PAR, ice-covers attenuating light, permanent low temperatures and the lack of a significant external input of inorganic nutrients and organic

carbon (Laybourn-Parry, 2002). Despite their generally oligotrophic to ultra-oligotrophic nutrient status, lakes are important sources of carbon in terrestrial Antarctic ecosystems given that they harbour many autotrophs, which can show substantial primary production (Laybourn-Parry, 2002). The vast majority of this primary production is mainly associated with the benthic microbial mats which are usually dominated by filamentous *Cyanobacteria*, most prominently of the order *Oscillatoriales* (Vincent et al., 2000; Cowan and Ah Tow, 2004; de los Ríos et al., 2015). In perennially ice covered lakes, anoxygenic phototrophy is probably important in the primary production (Laybourn-Parry and Wadham, 2014). Bacteria in the sediment, predominantly belonging to the phyla *Actinobacteria*, *Firmicutes*, *Bacteroidetes* and *Proteobacteria*, play significant roles in the remineralization of organic matter within aquatic ecosystems (Shivaji et al., 2011). However, in some systems, in order to deal with nutrient limitations, pathways have been adapted, short-circuiting biogeochemical cycles and preventing remineralization of these nutrients. Using a system-level approach, this was demonstrated to be the case for the nitrogen cycle in meromictic Ace Lake in the Vestfold Hills. During the transition from a marine to a closed, resource-limited stratified lake, particularly a *Chlorobi* (Green Sulfur Bacteria) species (*C*-Ace) became a dominant key species (Ng et al., 2010) and the reduction or expansion of genes and pathways were linked to this event (Lauro et al., 2011). Also unusual sulfur cycles have been observed, for example from Blood Falls (MDV) (Mikucki and Priscu, 2007) and Organic Lake (Yau et al., 2013).

1.5. Methods used in the study of bacterial diversity and communities in Antarctica

Early studies in the Antarctic focused on the physical and chemical characteristics and processes of the environment, and such studies have continued, particularly in extreme lake systems (e.g., Howard-Williams and Hawes, 2007). Biological research has generally also focused on lakes and involved mostly crustaceans, algal mats and larger protozoans

(Laybourn-Parry and Wadham, 2014). The diversity and distribution of terrestrial invertebrates have been investigated well, too, providing many insights into the controls and history of life on land in the Antarctic, as well as its interactions over the long term with glaciological and geological processes that have shaped the continent (Chong et al., 2015). However, despite their abundance and importance in Antarctic ecosystems, studies on Bacteria are fairly limited. Most research on Bacteria has been performed on the saline, marine derived lakes of the Vestfold, Larsemann and Bunger Hills of East Antarctica, lakes and soils in the Dry Valleys of Victoria Land, and to a lesser extent in the Maritime Antarctic and the Sub-Antarctic.

1.5.1. Culture-dependent studies

Isolation of Antarctic prokaryotes started as early as 1918, when McLean isolated Gram-positive cocci and Gram-positive and Gram-negative sporeforming and non-sporeforming rod-shaped bacteria from snow and ice and in association with frozen algae at Adélie Land near the South Pole (Straka and Stokes, 1960), and sporadic studies were performed with additional phenotypic tests. By 1960, the interest in potential psychrophilic growth and hence matter cycling at cold temperatures had grown (Straka and Stokes, 1960). Studies have been steadily increasing since the 1970s, and particularly since the 1980s has a considerable amount of Antarctic microbial studies been undertaken (Tindall, 2004). Early phylogenetic studies showed that isolated organisms displayed relatively large diversions from their closest relatives, suggesting that at least part of the contemporary prokaryotic diversity may have survived through multiple glacial cycles, and are not recent colonists (Franzmann and Dobson, 1993). From the late 1990s, there was a rise in isolates reported from diverse Antarctic habitats and many were indeed shown to be previously unknown species (Tindall, 2004).

It has long been recognized that there is a severe discrepancy between who (or what) can be grown in the lab compared to microscopic cell counts from environmental samples, an observation referred to as "the great plate count anomaly" (Staley and Konopka, 1985). Much effort has since been put into tailoring cultivation conditions to mimic the natural environment. For example, the often extreme oligotrophic *in situ* conditions have led to the use of carbon deprived growth media and prolonged incubation times, up to one year or longer (Brambilla et al., 2001; Bakermans et al., 2003, 2014; Stewart, 2012). Many isolates are readily grown at >10 °C (Peeters, Verleyen, et al., 2011), while relatively few are true psychrophilic organisms (Laybourn-Parry and Wadham, 2014). Attempts to isolate cryophilic organisms using sub-zero temperatures have been performed (Bakermans et al., 2003), and it has been demonstrated that they are metabolically active at these low temperatures in contrast to non-cryophilic organisms.

Cultivation has been used to target general (heterotrophic) bacterial diversity (e.g., Van Trappen et al., 2002; Barrientos-Díaz et al., 2008; Peeters et al., 2011) as well as specific heterotrophic groups, such as the phyla *Planctomycetes* and *Actinobacteria* (Brambilla et al., 2001) and Cyanobacteria (Taton et al., 2003). Such studies have resulted in both new species (e.g., Franzmann et al., 1987; Reddy et al., 2000, 2004; Van Trappen et al., 2002; Shivaji et al., 2005) and new genera (e.g., Bowman et al., 1998; Sheridan et al., 2003; Van Trappen et al., 2004; Sattley et al., 2008). Organisms that are easily grown include members of mainly *Proteobacteria*, *Firmicutes*, *Deinococcus-Thermus*, *Bacteroidetes* and *Actinobacteria* (Brambilla et al., 2001; Van Trappen et al., 2002; Peeters, Ertz, et al., 2011). Interest in isolates is also driven by biotechnological importance in application of psychrophilic and psychrotolerant enzymes and processes that can be active at lower temperatures (Margesin and Miteva, 2011).

Recent advances in growing previously uncultured species include co-culturing with other

bacteria, recreating the environment in the laboratory, and combining these approaches with micro-cultivation technology to increase throughput and access rare species (Stewart, 2012). The cultivation in the laboratory, combined with the now easily performed whole-genome sequencing, will lead to a better understanding of the physiology of these bacteria and their roles in ecology and natural product production (Kim et al., 2012; Stewart, 2012; Das et al., 2015).

1.5.2. Culture-independent methods

Despite their merits, cultivation approaches have only limited potential to comprehensively cover all diversity, and are labour intensive. It is estimated that only about one percent of the actual bacterial diversity can be cultivated with traditional methods (Hugenholtz, 2002), with a bias towards certain groups (Pham and Kim, 2012). In addition, it is not quantitative and does not permit any insights on community structures. This has been partly alleviated by the introduction of culture-independent techniques, which bypass the need to isolate organisms and directly use total DNA extracted from a sample. Much more studies throughout the Antarctic have used a culture-independent approach than a culture-based approach (Michaud et al., 2012). Clone library construction has been the most used culture-independent approach in Antarctica to date (e.g., Bowman et al., 2000; Brambilla et al., 2001; Aislabie et al., 2006, 2013), enabling the detection of uncultured organisms (Shivaji et al., 2011), or providing insights into functional diversity (Tahon et al., 2016a). In addition, fingerprinting or profiling methods have been applied. These techniques provide relatively fast and (in general) cheap ways to actually visualize community structure (van Dorst et al., 2014), either through band patterns on a gel, for instance with Denaturing Gradient Gel Electrophoresis (DGGE) (Muyzer et al., 1993; Pearce, 2005; Chong et al., 2009; Laybourn-Parry et al., 2013), or as peak patterns on a chromatogram, which is the case for Terminal Restriction Length

Polymorphism (t-RFLP) (Liu et al., 1997; Bowman, Rea, et al., 2000; Shivaji et al., 2011; Abramovich et al., 2012) and Automated Ribosomal Intergenic Spacer Analysis (ARISA) (Fisher and Triplett, 1999; Wood, Mountfort, et al., 2008; Wood, Rueckert, et al., 2008; Banks et al., 2009; Soo et al., 2009; Smith et al., 2010). Fingerprinting allows comparison between and within samples, revealing community dynamics (Grant and Ogilvie, 2004). It is generally assumed that only the dominant organisms (i.e., those which make up more than 0.5 to 1 % of the community) are picked up (Kan et al., 2006). These techniques do not allow a direct or unambiguous identification of the community members (van Dorst et al., 2014), since they do not give any sequence information, and unrelated organisms can result in the same signal or multiple signals can originate from a single organism (Popa et al., 2009). With DGGE, however, gel bands can be excised, purified and the DNA sequenced, which can then be compared to a database. Less frequently, the data output of both t-RFLP and ARISA can also be analysed further to obtain sequence information (van Dorst et al., 2014). Several studies have, however, combined multiple techniques, providing both insights in community structures and composition (Smith et al., 2006; Soo et al., 2009).

1.5.3. High-throughput Sequencing

Although traditional culture-independent approaches enabled a direct view on community structure or composition, the advent of high-throughput Next Generation Sequencing (NGS) again revolutionized microbial ecology. Pyrosequencing, although already developed in the late 1990s (Ronaghi et al., 1996, 1998), was not extensively commercialized until 2004 (Margulies et al., 2005). It was the first major NGS approach and, by shedding light on the so-called ‘rare biosphere’, revealed an even greater microbial diversity than previously expected (Sogin et al., 2006). However, it has since been clear that these approaches are also

susceptible to biases and errors (Huse et al., 2010; Kunin et al., 2010; Luo et al., 2012; Pinto and Raskin, 2012; Kennedy et al., 2014), and require sound analysis (Gomez-Alvarez et al., 2009; Miller et al., 2010; Edgar et al., 2011; Schloss et al., 2011). Nonetheless these methods provide unprecedented resolution and are increasingly used in molecular ecology studies, although their application on Antarctic samples is fairly limited up until now. Recently, technological improvements, resulting in increased read lengths and a higher output, have brought a new wave of sequencing platforms (third generation sequencing), such as Illumina's MiSeq. Because of the increased output compared to Sanger sequencing, the next and third generation sequencing platforms are jointly referred to as high-throughput sequencing approaches. Amplicon based sequencing, generally targeting ribosomal RNA, is the default approach, and has been applied to lake systems (Huang et al., 2014; Stephen D. J. Archer et al., 2015), soils (Charles K Lee et al., 2012; Pearce et al., 2012; I. Richter et al., 2014; Kim et al., 2015; Wang et al., 2015) and hypolithic environments (Makhalanyane et al., 2013), rhizospheres (Teixeira et al., 2010) and marine environments (Ghiglione and Murray, 2012). Recently, HTS has also been used in shotgun metagenomics surveys, looking at taxonomic and phylogenetic diversity, as well as functional potential (Lauro et al., 2011; Grzymski et al., 2012; Pearce et al., 2012).

1.6. Objectives

The main objective of this PhD-study was to explore the diversity and distribution of bacteria in diverse Antarctic soils and aquatic microbial mat samples using high-throughput sequencing approaches in order to expand our base line knowledge of bacterial diversity present in these extreme environments.

In a first study (Chapter 2), the bacterial community composition of two terrestrial and seven aquatic microbial mat samples from various regions in Antarctica was explored using the Roche 454 pyrosequencing, targeting the V1-V2 (forward) and V3-V2 (reverse) variable regions of the 16S rRNA gene. These samples had previously been used for the isolation of heterotrophic bacteria. As a second objective, the results of both approaches were compared.

While most microbiological studies focus on the lowland McMurdo Dry Valleys or other coastal sites, little is known about microbial life in inland regions and the environmental drivers structuring them. Therefore, in a second study (Chapter 3), the bacterial communities from 52 samples in the large vicinity of the Princess Elisabeth Station in the Sør Rondane Mountains (Dronning Maud Land, East Antarctica) were explored. The effect of different bedrock types (gneiss and granite) and the presence of macrobiota (mosses, lichens, microalgae) on the bacterial community structure and diversity were investigated. We also compared the results of both Illumina MiSeq sequencing and the ARISA genetic fingerprinting technique and related them to a range of environmental parameters and the geographic distance between samples.

Accumulating evidence shows that biogeographical patterns exist for microorganisms. However, such patterns have not been unambiguously established in the Antarctic, with many sampling campaigns being limited to a few regions and covering a small spatial scale, and

employing different methods. In a last study (Chapter 4), biogeographical patterns in microbial mat bacterial communities were explored in a large-scale survey, using a uniform approach. Herefore, 138 lakes along a broad gradient of limnological conditions from eight different regions in Antarctica and two Sub-Antarctic islands were selected, and sequenced using Illumina MiSeq platform.

This PhD study is part of the Climate Change and Antarctic Microbial BIOdiversity (CCAMBIO) project, funded by the Belgian Science Policy Office (BelSPO). CCAMBIO is coordinated by the Centre for Protein Engineering (University of Liège), who focus on *Cyanobacteria*, in collaboration with the Laboratory of Protistology and Aquatic Ecology (Ghent University), focusing on the eukaryotic microorganisms, and the Laboratory of Microbiology, studying the general bacterial diversity. It is the successor to the Antarctic Microbial BIOdiversity (AMBIO) project.

Chapter 2. Bacterial diversity assessment in Antarctic terrestrial and aquatic microbial mats: a comparison between bidirectional pyrosequencing and cultivation

Redrafted from:

Tytgat, B., Verleyen, E., Obbels, D., Peeters, K., De Wever, A., D'hondt, S., De Meyer, T., Van Criekinge, W., Vyverman, W. and Willems, A. (2014). Bacterial diversity assessment in Antarctic soils and lake sediments: a comparison between bidirectional pyrosequencing and cultivation. *PlosOne*: vol. 9, issue 6. 10.1371/journal.pone.0097564

Authors' contributions:

BT, AW, EV, ADW, KP and WV designed the study. ADW, DO, SD, TDM and WVC prepared the pyrosequencing. KP performed the isolation campaign. BT, DO and EV analysed the data. BT, EV, TDM, WVC, WV and AW wrote the manuscript

2.1. Summary

The application of high-throughput sequencing of the 16S rRNA gene has increased the size of microbial diversity datasets by several orders of magnitude, providing improved access to the rare biosphere compared with cultivation-based approaches and more established cultivation-independent techniques. By contrast, cultivation-based approaches allow the retrieval of both common and uncommon bacteria that can grow in the conditions used and provide access to strains for biotechnological applications. We performed bidirectional pyrosequencing of the bacterial 16S rRNA gene diversity in two terrestrial and seven aquatic Antarctic microbial mat samples previously studied by heterotrophic cultivation. While, not unexpectedly, 77.5 % of genera recovered by pyrosequencing were not among the isolates, 25.6 % of the genera picked up by cultivation were not detected by pyrosequencing. To allow comparison between both techniques, we focused on the five phyla (*Proteobacteria*, *Actinobacteria*, *Bacteroidetes, Firmicutes* and *Deinococcus-Thermus*) recovered by heterotrophic cultivation. Four of these phyla were among the most abundantly recovered by pyrosequencing. Strikingly, there was relatively little overlap between cultivation and the forward and reverse pyrosequencing-based datasets at the genus (17.1-22.2 %) and OTU (3.5-3.6 %) level (defined on a 97 % similarity cut-off level). Comparison of the V1-V2 and V3-V2 datasets of the 16S rRNA gene revealed remarkable differences in number of OTUs and genera recovered. The forward dataset missed 33 % of the genera from the reverse dataset despite comprising 50 % more OTUs, while the reverse dataset did not contain 40 % of the genera of the forward dataset. Similar observations were evident when comparing the forward and reverse cultivation datasets. Our results indicate that the region under consideration can have a large impact on perceived diversity, and should be considered when comparing different datasets. Finally, a high number of OTUs could not be classified using the RDP reference database, suggesting the presence of a large amount of novel diversity.

2.2. Introduction

With its severe physical, chemical, and climatic conditions (Cowan and Ah Tow, 2004), Antarctica is characterized by harsh environmental settings and hosts communities of well-adapted microbiota that are capable of withstanding selective pressures, such as high UV-radiation, drought, light limitation and extremely low temperatures. These adaptations may therefore be potentially of biotechnological and economical value (Cavicchioli et al., 2002; Loperena et al., 2012). Until now, studies have mainly used culturing approaches (Van Trappen et al., 2002; Peeters, Verleyen, et al., 2011) and a number of culture-independent techniques such as Denaturing Gradient Gel Electrophoresis (DGGE) (Ghiglione and Murray, 2012), Terminal Restriction Fragment Length Polymorphism (t-RFLP) (Pointing et al., 2009b; Shivaji et al., 2011), Automated Ribosomal Intergenic Spacer Analysis (ARISA) (Soo et al., 2009) and clone libraries (Aislabie et al., 2006; Newsham et al., 2010; Pearce et al., 2010; Shivaji et al., 2011) to shed light on Antarctic bacterial diversity. These studies reported taxa that are new to science (Brambilla et al., 2001; Van Trappen et al., 2002; Peeters, Verleyen, et al., 2011) and/or revealed that – as in other regions and environments (Kautz et al., 2013) – Antarctic microbial diversity is much larger than previously thought.

Whereas Next Generation Sequencing (NGS) techniques have now found their way to nearly every environment, ranging from the deep sea (Huber et al., 2007) to tropical forest soils (Leff et al., 2011) and the human microbiome (Huttenhower et al., 2012), the Antarctic region remains relatively underrepresented in these microbial diversity studies. This is surprising, given the fact that the diversity reported with NGS is orders of magnitude higher than that recovered with traditional culturing and Sanger sequencing, and at least one order of magnitude higher than recovered from large clone libraries (Pedrós-Alió, 2012). More recently NGS has been used to study Antarctic samples, including McMurdo Dry Valley soils

(Charles K Lee et al., 2012; Van Horn et al., 2013), soils from Alexander Island (Pearce et al., 2012), rhizosphere bacteria of the only two vascular plants in the Antarctic Peninsula (Teixeira et al., 2010), a study of community turnover due to global warming (Yergeau et al., 2012), a survey of cyanobacterial diversity in microbial mats (Varin et al., 2012) and a comparison of seasonal variation in coastal marine bacterioplankton (Ghiglione and Murray, 2012). The relative paucity of Antarctic studies is largely due to the remoteness and vastness of the continent, the harsh environmental conditions and the costs associated with expeditions. Yet, exactly these limitations have kept the environment relatively pristine, thus providing excellent conditions to investigate several questions of particular interest to microbiologists such as to which extent historical processes shape microbial biogeography patterns and the degree of endemism. Moreover, polar regions with their uniquely adapted microbiota are particularly prone to the impact of global warming (Thompson and Solomon, 2002; Walther et al., 2002; Kirchman et al., 2009; Hodgson, 2011) and microbial diversity data are therefore urgently needed as baseline data for tracking this impact.

Microbial communities typically consist of few high-abundant taxa, with the majority of taxa belonging to the so called rare biosphere (Sogin et al., 2006, 2010; Pedrós-Alió, 2012). Although it was shown that cultivation is able to pick up some of these rare community members (Shade et al., 2012), it is generally thought that only through the deep sequencing that NGS offers, this vast diversity can be detected (Charles K. Lee et al., 2012; Pedrós-Alió, 2012). In turn, this also implies that cultured strains are generally expected to be recovered by pyrosequencing. Here we aimed to test this hypothesis by comparing the diversity of heterotrophic bacterial groups previously recovered from Antarctic microbial mat samples by cultivation with the diversity of the corresponding groups as revealed by 454 pyrosequencing. An additional objective was to assess the impact of the region of the 16S rRNA gene on the diversity data obtained. This was done by comparing forward and reverse pyrosequencing

datasets and contrasted with a comparison of forward and reverse data from the cultured strains, where no effects of the pyrosequencing process could be at work.

2.3. Material and Methods

Samples used

Details of the study sites have been described previously (Peeters and Willems, 2011; Peeters, Ertz, et al., 2011; Peeters, Hodgson, et al., 2011). Briefly, two terrestrial and seven limnetic microbial mat samples were collected aseptically during different field campaigns in December/January 2003 (PQ1, TM2 and TM4) and in January 2007 (BB50, BB115, LA3, SK5, WO10 and SO6). One sample (PQ1) was collected on Pourquoi-Pas Island off the west coast of Graham Land (Antarctic Peninsula). All other samples were collected from Eastern Antarctic habitats. The two terrestrial microbial mat samples (BB50 and BB115) were taken near the Utsteinen nunatak in the Sør Rondane Mountains (Dronning Maud Land). Three samples were from Lützow-Holm Bay (Dronning Maud Land), namely from a small saline lake in Langhovde (LA3), from Naka-Tempyo Lake (SK5) in Skarvsnes, and from a small saline pond (WO10) in West Ongul Island. One sample (SO6) was taken from Lake Melkoye (unofficial name) in Schirmacher Oasis (Dronning Maud Land). The two remaining samples were collected in the Transantarctic Mountains. Sample TM2 was taken from Forlidas Pond (Dufek Massif, Pensacola Mountains), while sample TM4 was taken from Lundström Lake (Shackleton Range). All samples were kept frozen during transport and stored at -20 °C.

Processing of 16S rRNA gene sequences of cultures

The cultured heterotrophic bacterial diversity of these samples was reported earlier (Peeters and Willems, 2011; Peeters, Ertz, et al., 2011; Peeters, Hodgson, et al., 2011; Peeters, Verleyen, et al., 2011). From these, we selected 1,666 high quality sequences for comparison with pyrosequencing. To allow this comparison, the sequences from bacterial cultures were

aligned to the Silva reference database (Pruesse et al., 2007), and trimmed so as to cover the alignment of the sequences obtained using pyrosequencing (see below). They were further processed together with both forward and reverse pyrosequencing datasets.

Pyrosequencing

To allow direct comparison, DNA was extracted from the same frozen samples previously used for the cultivation experiments using 5 g per sample. Extracellular DNA was first removed as described by Corinaldesi et al. (2005), by adding 2.5 g (wet weight) of sediment to 7.5 ml of 0.1 M sodium phosphate buffer (pH 8.0) with 0.5 g of acid-washed polyvinylpolypyrrolidone. Samples were homogenized with a horizontal shaker at a low speed (150 horizontal shakes per minute) for three 1-min cycles, with 1 min of cooling on ice between cycles. Subsequently, SDS (final concentration, 0.1 %) was added, and the samples were shaken again for 10 s. Then the samples were chilled on ice and centrifuged at 500 × g for 10 min at 4 °C, and the supernatants were removed. The sediment pellets were washed two more times by adding 7.5 ml of 0.1 M sodium phosphate buffer (pH 8.0) and centrifuged as described above (but in this case there was no addition of SDS). DNA extraction was subsequently performed according to Zwart et al. (1998). To lyse the cells, 0.5 g of zirconium beads (0.1 mm diameter), 0.5 ml TE buffer (10 mM Tris, pH 7.6, 1 mM EDTA) and 0.5 ml buffered phenol (pH 7 to 8) were added to the tubes containing the cells and the tubes were vigorously shaken (5000 rpm) on a Mini Bead-beater (Biospec Products, Bartlesville, OK, USA) for two min with intermittent cooling on ice. The tubes were then centrifuged for 5 min at 10,000 × g and the upper (aqueous) phase was collected and extracted twice with phenol-chloroform-isoamylalcohol (25:24:1). The DNA was then precipitated by adding one tenth volume of 3 M sodium acetate (pH 5) and two volumes of 96% (v/v) ethanol and centrifuging for 30 min at 14,000 × g. Subsequently, the DNA was dissolved in water and purified on a Wizard column (Promega, Madison, WI, USA) according to the manufacturer's

recommendations.

Sequencing of the 16S rRNA V1-V3 regions was performed using forward primer pA (AGAGTTTGATCCTGGCTCAG 8-27) (Edwards et al., 1989) and reverse primer BKL1 (GTATTACCGCGGCTGCTGGCA 536-516). Because it proved impossible to concatenate the complementary reads due to insufficient overlap, the forward and reverse sequences were analyzed separately. The forward reads hence cover the complete V1 and V2 regions, whereas the reverse reads cover the V3 and part of the V2 region for the longest sequences (Chakravorty et al., 2007).

Multiplexing was done with barcodes proposed by Parameswaran et al. (2007). Each PCR mixture contained 1-2 µl of template DNA, 2 µl of fusion primers and barcodes (10 µM), 2.5 µl dNTPs (10 mM), 1.5 µl of 10x buffer, 0.25 µl of 5 U/µl FastStart High Fidelity Polymerase (Roche) and was adjusted to a final volume of 25 µl with sterile HPLC-water. PCR cycling included 3 min at 94 °C, followed by 35 cycles of 94 °C for 30 s, 55 °C for 60 s and 72 °C for 90 s and finally 8 min at 72 °C. PCR products were purified using a High Pure PCR Product Purification Kit (Roche). Finally, pyrosequencing was performed on a Roche 454 GS FLX Titanium machine at NXTGNT (Ghent, Belgium) after quality control of the DNA with a Qubit 2.0 Fluorometer (Life Technologies) and a Bioanalyzer (Agilent Technologies).

Raw sequences are available from the NCBI Sequence Read Archive under accession numbers SRR1146576 and SRR1146579.

Processing of pyrosequences

The obtained reads were processed using Mothur (Schloss et al., 2009) version 1.27.0, generally following Schloss et al. (Schloss et al., 2011) and the Mothur SOP (http://www.mothur.org/wiki/454_SOP; version of 6 November 2012). The data were denoized using Mothur's PyroNoise (Quince et al., 2009) implementation with 450 flows as

the minimal flow length and trimming of the longer sequences to this length (Schloss et al., 2011). Overall, the minimal required sequence length was set at 200 nucleotides (nt). To avoid poor sequence quality, no ambiguous bases (N) were allowed (Huse et al., 2007) and sequences with homopolymers longer than 8 nt were culled, as it is known that long homopolymers are problematic for 454 pyrosequencing (Datta et al., 2010; Gilles et al., 2011; Schloss et al., 2011). The sequences were aligned using Mothur's alignment command, based on the GreenGenes NAST aligner (DeSantis, Hugenholtz, Keller, et al., 2006) with default parameters and the Silva reference database (Pruesse et al., 2007), which takes into account the secondary structure of the 16S SSU rRNA. The starting and ending positions of the alignment were checked to ensure that sequences were overlapping the same alignment space. Sequences not starting at the correct position or ending before 95 % of all the sequences were removed from the analysis. To increase computational speed and decrease data size, duplicate (identical) sequences were temporarily removed using the unique.seqs command. Further correction for erroneous base calls was done using single linkage preclustering according to Huse et al. (2010). Next, we used Uchime (Edgar et al., 2011) with default parameters for intra-sample *de novo* chimera checking. Positively identified chimeric sequences were removed from further analyses.

Sequence identification and OTU clustering

Sequences were identified using Mothur's implementation of the RDP classifier (Wang et al., 2007) by means of the modified RDP training-set release 9 (available at http://www.mothur.org/wiki/RDP_reference_files) at an 80 % bootstrap value. The RDP database was chosen so that a comparison with the original cultivation data was possible, despite its known limitations because of its small size (Newton and Roeselers, 2012; Werner

et al., 2012), possibly overestimating the number of unclassified OTUs. This training set too was first aligned and trimmed to the alignment space of the query sequences, increasing confidence values and reducing the number of unclassified sequences (Werner et al., 2012). Non-cyanobacterial "chloroplast" sequences were removed from the dataset. Distances were calculated (dist.seqs command, default settings), after which the sequences were clustered using the average neighbor joining algorithm to generate OTUs at a 97 % cutoff level (Schloss et al., 2011).

SIMPROF analysis

In order to compare the community composition obtained using culturing versus pyrosequencing a SIMPROF analysis (Clarke et al., 2008) was performed using Primer 6. SIMPROF is a permutation-based procedure that ranks the pairwise similarities in each group and tests the null hypothesis that samples were all drawn from the same species assemblage. Because the number of sequences is consistently higher in the pyrosequencing dataset, we standardized the number of sequences in each sample to the lowest number of sequences obtained in all of the samples (i.e., 119 forward and 116 reverse sequences in sample LA3). To achieve this, we randomly sampled this number of sequences from each sample with replacement. This procedure was done 5 times, which resulted in 5 subsets for each sample. First, a Jaccard similarity matrix was constructed and subsequently used to undertake a group-average cluster analysis. Second, to ascertain the level of structure present in the groups formed by each dendrogram, a SIMPROF test with 10,000 simulations and the stopping rule specified at the 5 % significance level was run. This was done for both forward and reverse datasets.

2.4. Results

Sequence data of bacterial isolates

Of the initial 1,666 sequences, 1,578 remained after the forward processing together with the pyrosequences. This was mainly due to the removal of sequences that did not match the correct starting or ending positions of the alignment space. A total of 342 OTUs in 76 genera from five different phyla were obtained (Figure 2.1). Most of the OTUs belonged to the phyla *Bacteroidetes* and *Proteobacteria*, with 107 and 106 members respectively. *Actinobacteria* was the third best represented phylum with 78 OTUs, followed by *Firmicutes* and *Deinococcus-Thermus* with 31 and 20 OTUs respectively.

The initial 1,666 sequences were also subjected to the reverse processing pipeline. In contrast to the 1,578 forward sequences, this yielded only 1,519 sequences divided over 214 OTUs in 61 genera. The relative proportion of the phyla did not differ drastically when processed through the forward or reverse pipeline (Figures 2.1 and 2.2), although only 51 genera were shared between the forward and reverse dataset of the isolates. In total, we identified 86 genera for the combined processed cultivation results, while some sequences remained unclassified. Of these 86 genera, 20 (23 % of cultivated genera) were not picked up by pyrosequencing.

Heatmaps showing the distribution of the most frequently recovered OTUs based on the forward (Figure S2.1) and reverse (Figure S2.2) cultivation sequences, revealed that many of these OTUs were shared between samples.

Pyrosequencing data

Forward dataset

After processing the forward pyrosequencing data, 23,510 high quality sequences were left (on average 2612 ± 829 per sample); they were on average 243 ± 14 nt long. The chimera content per sample in the forward dataset ranged from 0.1 % (TM2) to 5.8 % (SK5) of sequences (Table S2.1). For eight samples, in the non-redundant dataset (i.e., dataset filtered for duplicate sequences), the percentage of chimeras was higher than when considering the complete dataset, indicating that many chimeras were singletons or low-abundant sequences.

We observed 2940 OTUs of which 947 remained unclassified at the phylum level (represented by 7659 sequences) and 2066 (15,271 sequences) at the genus level. Per sample, the number of OTUs unclassified at the phylum level varied between 40 (TM4) and 274 (WO10). The identified OTUs belonged to 220 genera in 19 phyla (Tables S2.2 and S2.3 respectively)[3,4]. *Proteobacteria*, *Cyanobacteria*, *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, *Acidobacteria* and *Planctomycetes* were present in every sample (Figure 2.1), although relative number and OTU richness could differ drastically. Cyanobacteria were well represented in most samples, but less so in SK5 and BB115. *Deinococcus-Thermus* was relatively well recovered and showed a high richness in the terrestrial samples (BB50 and BB115).

A total of 2693 (84.9 %) of the OTUs were restricted to one sample (Figure 2.3), and 1464 (46.2 %) were effectively singletons (i.e., represented by only one sequence).

[3] Table S2.2 online available at:
http://journals.plos.org/plosone/article/asset?unique&id=info:doi/10.1371/journal.pone.0097564.s010

[4] Table S2.3 online available at:
http://journals.plos.org/plosone/article/asset?unique&id=info:doi/10.1371/journal.pone.0097564.s011

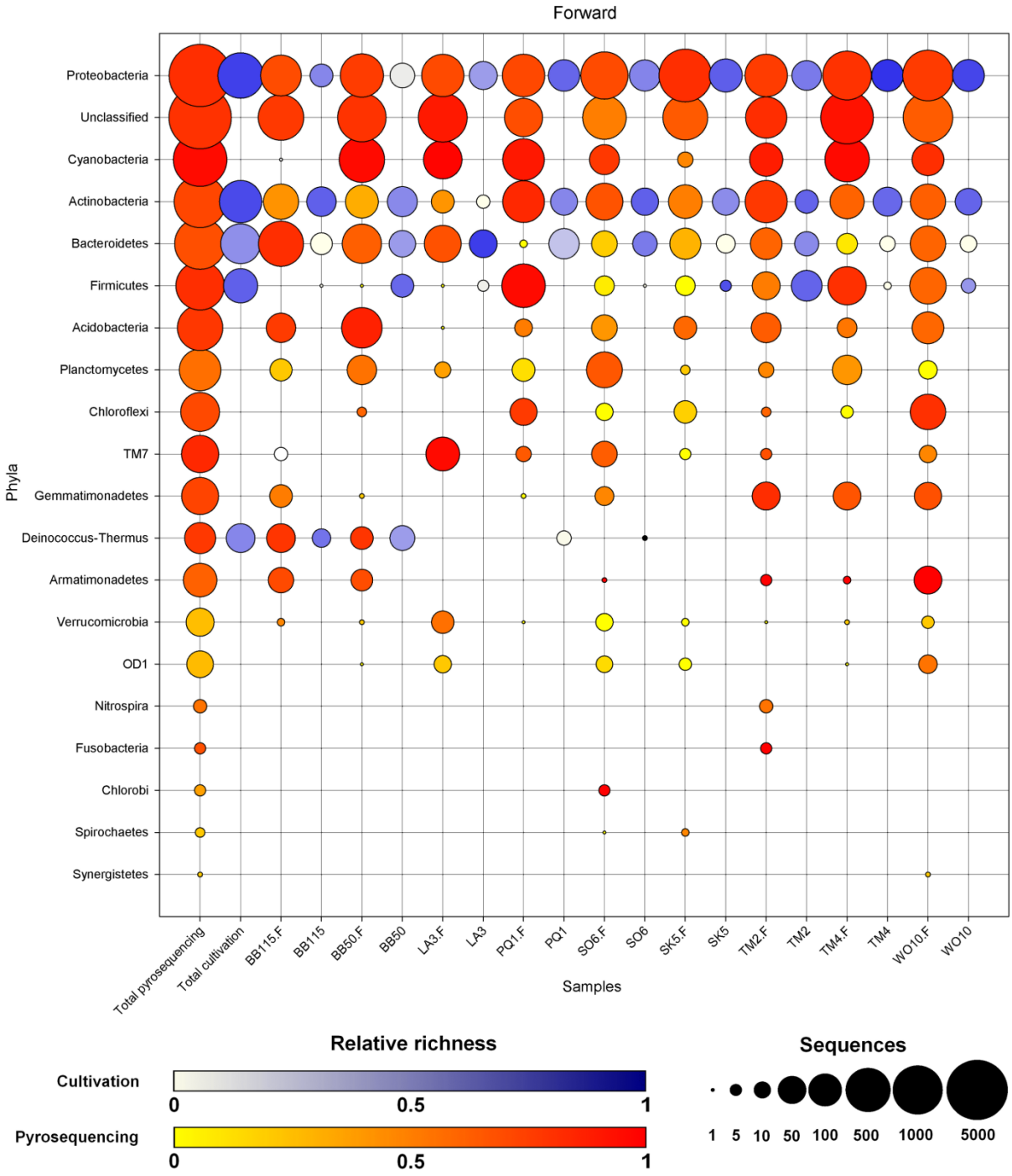

Figure 2.1. Overview of the distribution of the phyla per sample for the forward sequencing dataset. Circle area is a $\log_2$ transformation of the number of sequences ([$\log_2$(N)*5/PI], with N the number of sequences in that phylum). Color intensity reflects the number of OTUs per phylum (total OTUs/total sequences), with a darker hue indicating a higher relative richness. The first two columns show the total number of sequences and diversity of each phylum for pyrosequencing and cultivation separately. The phyla are ordered according to decreasing total number of sequences. The yellow to red scale shows pyrosequencing data, the blue-purple scale the cultivation data.

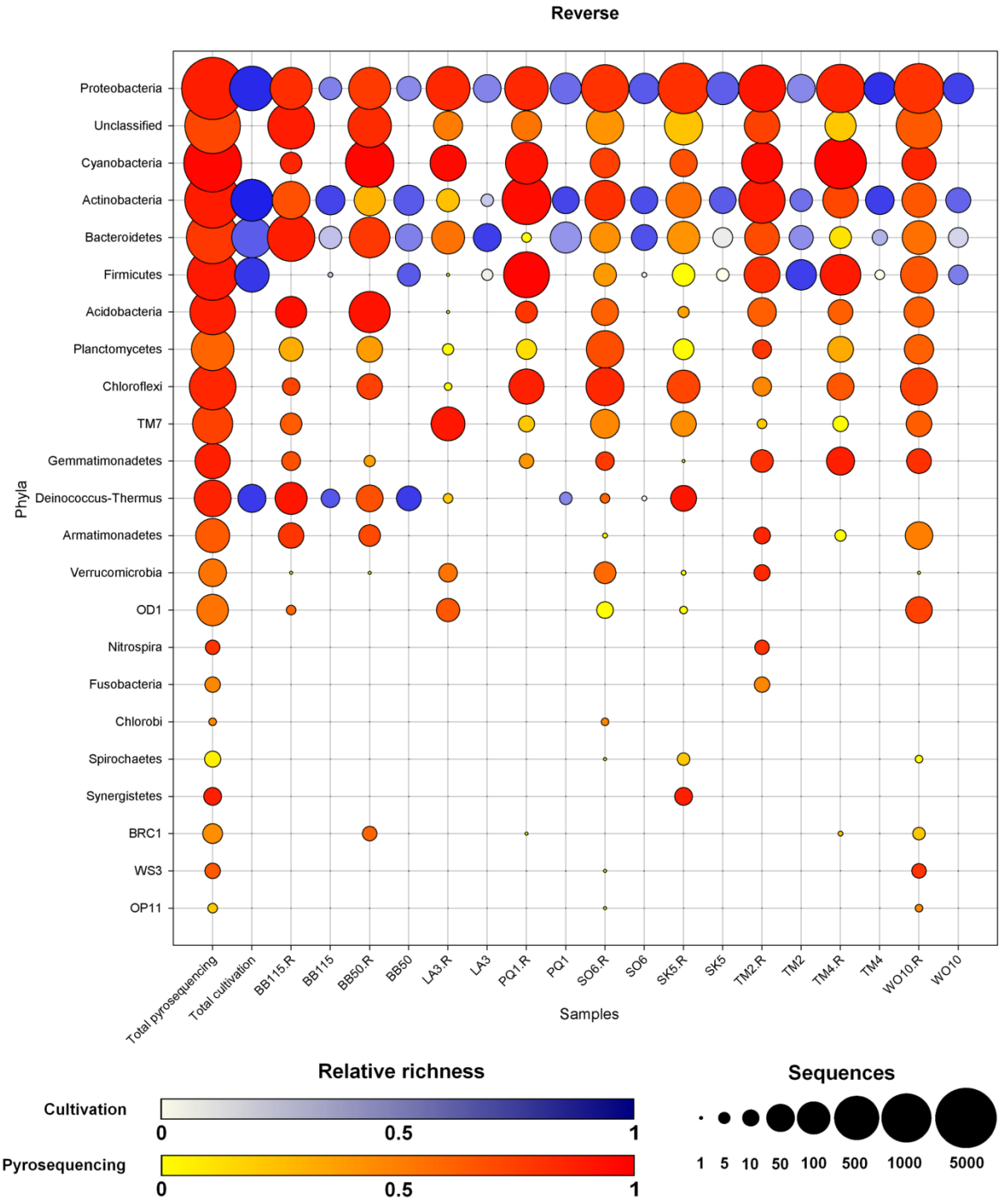

Figure 2.2. Overview of the distribution of the phyla per sample for the reverse sequencing dataset. Circle area is a $\log_2$ transformation of the number of sequences ([$\log_2$(N)*5/PI], with N the number of sequences in that phylum). Color intensity is an approximation for the number of OTUs per sequence (total OTUs/total sequences). The first two columns show the total number of sequences and diversity of each phylum for pyrosequencing and cultivation separately. The order of the phyla is as in Figure 1 and additional phyla were added at the bottom. The yellow to red scale shows pyrosequencing data, the blue-purple scale the cultivation data.

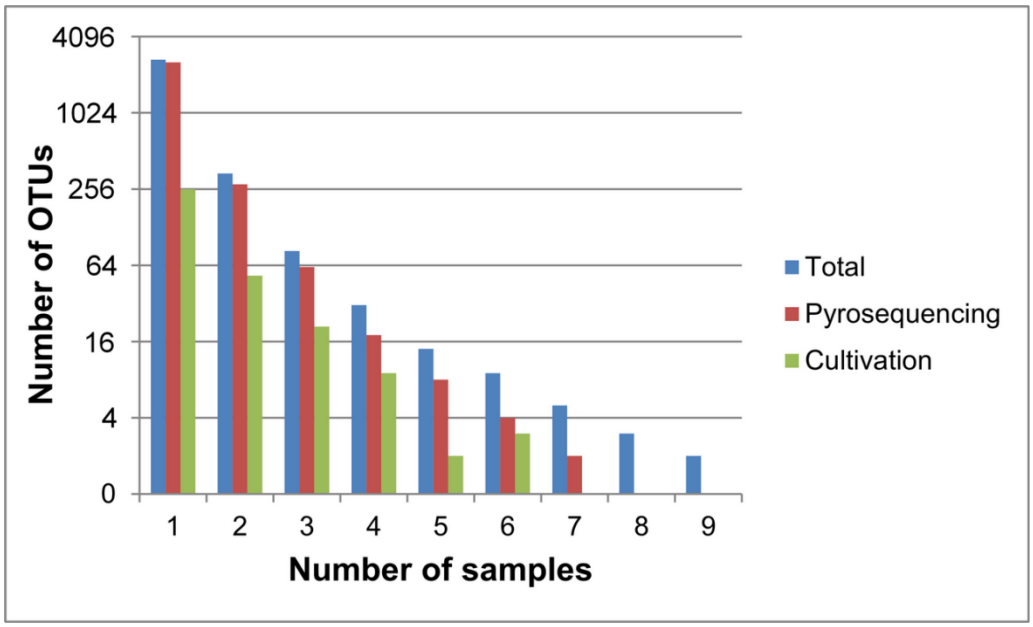

Figure 2.3. Bar chart illustrating the number of OTUs picked up from one or more samples for the forward dataset. The number of OTUs is $\log_2$ transformed. Blue bars, total sequences (pyrosequences plus cultivated sequences); red bars, pyrosequences only; green bars, cultivation sequences only.

The most abundant OTU (OTU3056) was represented by 2216 sequences, nearly three times as many as the second most abundant OTU (OTU0858, 871 sequences), and was found in six out of the nine samples (BB115, BB50, PQ1, TM2, TM4 and WO10). It was not picked up through cultivation and was not identified using our RDP training set. A separate blast against the GreenGenes database (DeSantis, Hugenholtz, Larsen, et al., 2006), however, revealed that it was identical to *Phormidium autumnale* str. Arct-Ph5 (*Cyanobacteria*, a group not targeted by the cultivation experiments). None of the OTUs was found in every sample through pyrosequencing in the forward dataset. One OTU (OTU2885; *Rhizobacter*, *Gammaproteobacteria*) was found in seven samples. Five OTUs were recovered by pyrosequencing from six samples, including the aforementioned cyanobacterial OTU3056, a *Polaromonas* (OTU2491, which was also cultured) and a *Herbaspirillum* species (both *Betaproteobacteria*), and a *Methylobacterium* species (*Alphaproteobacteria*) and finally OTU2399, identified as *Brevundimonas* sp. (*Alphaproteobacteria*), which was actually

retrieved from every sample (i.e., it was recovered through either cultivation, pyrosequencing or in some samples by both). All other OTUs were only recovered from five samples or less.

The combined number of OTUs from forward pyrosequencing and cultivation was 3172 (totaling 25,088 sequences). Only 110 OTUs were shared between both approaches and 232 were restricted to the cultivation data. A heatmap (Figure S2.3) showing the distribution of the most frequently recovered pyrotag OTUs, revealed that few of these OTUs were shared between samples. In fact, most of these high-abundant OTUs were merely recovered from one or two samples. The SIMPROF analysis revealed that the community structure in all samples assessed using pyrosequencing is significantly different from that analyzed using culturing (Figure S2.4). Not unexpectedly, given cultivation bias, the similarity between samples analyzed with the culturing approach is higher. However, these observations were consistent when taking into account only the five phyla that were recovered by both approaches (data not shown).

Reverse dataset

Reverse pyrosequencing starting from the end of the V3 region resulted in 22,778 high quality sequences after processing. The chimera content was generally higher than for the forward pyrotags for all samples (Table S2.1). Particularly in sample SK5, up to 43.4 % of the non-redundant sequences were identified as chimeras by Uchime, resulting in the removal of 23 % of all sequences in that sample. Also for sample PQ1 23 % of all sequences were removed, while only 19.6 % of the unique sequences were flagged as chimeras, indicating a substantial proportion of chimeras in this sample. We obtained only 1983 OTUs overall, of which 485 remained unclassified at phylum level (2776 sequences) while the rest belonged to 22 phyla (Figure 2.2, Table S2.3). We were able to identify 197 genera in the reverse dataset (Table S2.2). The taxonomy at genus level remained unresolved for 1376 OTUs (12,295 sequences).

Although considerably fewer OTUs were observed in the reverse dataset, the distribution over phyla were similar to these observed for the forward pyrosequences (Figure 2.2, Table S2.3). The number of sequences unclassified at phylum level (485 OTUs, 2776 sequences) was much smaller than in the forward sequencing (947 OTUs, 7,659 sequences) and represented 24 % versus 32 % of the OTUs, respectively. Compared to the forward dataset, *Deinococcus-Thermus* was additionally picked up from samples LA3, SO6 and especially SK5 (Figure 2.2). Also *Cyanobacteria* and *Chloroflexi* were generally more abundantly picked up by the reverse sequencing, and additionally, three extra bacterial phyla were recovered: WS3, OP11 and BRC1. Phylum BRC1 was obtained from four different samples (BB50, PQ1, TM4 and WO10) with six OTUs in total; WRC3 was represented by two OTUs, one from SO6 and a second one from WO10; OP11 was also found in these two latter samples. The number of singleton OTUs was lower for the reverse dataset: 476 (24 %) here vs. 897 (31 %) in the forward dataset. This discrepancy equals 44 % of the difference in the total number of OTUs obtained between both datasets (1983 in the reverse dataset compared to 2940).

Heatmaps showing the distribution of the OTUs most abundantly recovered in the reverse pyrosequencing data (Figure S2.5) and in the reverse cultivation dataset (Figure S2.2) reveal generally similar trends as for the forward sequencing (Figures S2.1 and S2.3). However, nine OTUs (1942, 1956, 1959, 2036, 2043, 2044, 2064, 2109 and 2115) in the high-abundant reverse pyrosequencing selection were also found in the cultivated dataset, which is considerably more than for the forward dataset. Especially OTU2109 (*Sphingomonadaceae* sp., *Alphaproteobacteria*) was recovered well through cultivation (not found in sample TM4), and pyrosequencing (not found in sample BB115). OTU1849 (*Methylobacterium*, *Alphaproteobacteria*) was recovered from all pyrosequencing samples. Four unclassified OTUs were recovered from eight samples (three alphaproteobacteria and one actinobacterium). The most abundant OTU (OTU1804) with 1,226 sequences was found in

five samples. It was classified as an unknown cyanobacterial order by the RDP training set. Again, a blast against the Greengenes database resulted in *P. autumnale* (strains Ant-Ph68 and Arct-Ph5, both with an identity score of 100). Similar to the forward dataset, both techniques resulted in significantly different clusters and the variability between the different samples is higher in the datasets obtained through pyrosequencing (Figure S2.6).

2.5. Discussion

Comparison of forward and reverse datasets

Two terrestrial and seven aquatic Antarctic microbial mat samples were subjected to bidirectional pyrosequencing of the V1 to V3 variable region of the 16S rRNA gene. After processing, the forward dataset spanned the V1 and V2 variable regions, while the reverse dataset covered the V3 and part of the V2 variable regions. The comparison of bidirectional sequencing revealed large differences in the number of OTUs recovered, although the number of sequences and genera was generally comparable. More in particular, the number of OTUs was about 50 % higher for the forward dataset compared to the reverse dataset. This is in part likely due to the V1 region being more variable than the more conserved V3 region (Yu and Morrison, 2004; Youssef et al., 2009; Jeraldo et al., 2011). Hence, the traditionally used cut-off values (e.g., 95 % as a proxy for genus level, or 97 % for species level) which have proven to be insufficient or inadequate for all taxa (Schloss and Handelsman, 2005), might additionally require modification for different regions of the 16S rRNA gene. Highly variable regions such as V1 could be clustered using lower values (for example 97 %) than more conserved regions (e.g., V3 or V6), which might require a higher (e.g., 99 %) identity cut-off. These considerations should be taken into account when selecting the region to analyze, but also when comparing studies and diversity data based on different variable regions

(Engelbrektson et al., 2010). Not only did the number of OTUs differ between both regions, identification was affected too. For example, although the number of genera identified from the forward and the reverse dataset was broadly similar (220 vs. 197), only 132 or 67 % of the genera identified from the reverse dataset were also present in the forward dataset, corresponding to 60 % of the genera in the forward dataset. The combined number of genera based on the RDP training set was 285. Similarly, for the Sanger sequences of the cultures, comparison of forward and reverse trimmed dataset revealed 76 and 61 genera respectively, of which 51 were in common. As pyrosequencing artifacts cannot have been introduced in the cultivation dataset, these differences highlight the impact of the variable zones covered on the outcome of the genus identifications. With the continuous development of NGS techniques, the significance of this problem can be expected to reduce with increasing read length.

Another striking difference between the sequencing directions was that the number of chimeras was higher in the reverse dataset (Table S2.1). This is probably also due to the differences in variability of the regions targeted; the more conserved V3 region might be more likely to function as a template for annealing than V1, especially between closely related taxa (Gomez-Alvarez et al., 2009; Haas et al., 2011). Furthermore, not only do PCR conditions (such as extension times and the number of PCR cycles) or conserved regions affect chimera formation (Acinas et al., 2004; Haas et al., 2011; Schloss et al., 2011), it has been shown that certain positions in the 16S rRNA gene are more prone to chimera formation (Gomez-Alvarez et al., 2009). This implies that chimeras are not necessarily restricted to low-abundant sequences, questioning the removal of only OTUs with a low abundance, a common practice to reduce artifacts.

Contrast between diversity data from pyrosequencing and cultivation

The comparison of the bacterial diversity estimate obtained by bidirectional 454 pyrosequencing with the results from previous cultivation studies (Peeters, Ertz, et al., 2011;

Peeters, Hodgson, et al., 2011; Peeters, Verleyen, et al., 2011) unsurprisingly confirmed that pyrosequencing results in a higher diversity (in total 22 phyla, 285 genera) than obtained through culturing (5 phyla, 86 genera). Indeed, we observed a striking and significant difference in taxonomic composition and abundance of groups recovered using both methods, with communities standardized to the lowest number of sequences (Figures S2.4 and S2.6). This likely results from the obvious bias related to the specific cultivation conditions used, which were set to target mostly heterotrophic, aerobic and psychrophilic or psychrotolerant bacteria (Peeters, Verleyen, et al., 2011). Some of the phyla that were detected by pyrosequencing but not picked up through cultivation included groups that were not targeted such as anaerobes (e.g., *Clostridium* which was frequently recovered in sample PQ1), phototrophic *Cyanobacteria* and *Chloroflexi*, or groups for which cultivation is not yet optimized and that have very few or even no cultured representatives (e.g., *Acidobacteria*, *Planctomycetes*, *Verrucomicrobia*, *Armatimonadetes*, TM7; see Table S2.3). Given that only heterotrophic bacteria had been targeted in the isolation campaigns and a limited set of cultivation conditions was tested, a comparison with pyrosequencing is only possible to a very limited extent. We tried to take this into account by further focusing this part of the discussion on the OTUs and named genera of the five phyla picked up by both techniques (*Actinobacteria*, *Bacteroidetes*, *Deinococcus-Thermus*, *Firmicutes* and *Proteobacteria*). This restricted comparison confirmed the general observation that pyrosequencing can detect more diversity at all taxonomic levels. Nevertheless, particularly at lower taxonomic and phylogenetic levels (OTUs, genera), we found extremely little overlap in the diversity between both datasets. For example, in the forward sequencing datasets, of the 342 OTUs recovered using culturing, 232 (67.8 %) were not picked-up by pyrosequencing. For these five phyla, a total of 204 genera were identified, of which 51 were in common, 131 were unique for pyrosequencing and 22 unique for cultivation. Thus about 30 % of cultured genera were

not detected in our pyrosequencing data (e.g., the *Firmicutes* genus *Paenibacillus*; see Table S2.2). Reverse sequencing showed generally analogous results.

In addition to the above mentioned cultivation bias, at least three other non-mutually exclusive processes might underlie the significant differences between the cultivation and pyrosequencing datasets. Firstly, manual picking of individual colonies for further characterization in culture-based approaches introduces an additional bias. The sheer quantity of isolates makes it nearly impossible to select and cultivate every colony separately, especially when the number of samples is high. Phenotypic (morphological) selection may thus lead to an underestimation of the genotypic diversity, because macroscopically identical colonies might in fact represent different OTUs, whether closely related or not. Secondly, the failure of pyrosequencing to detect the majority of the cultured organisms could indicate that our sequencing depth was not large enough (Figures S2.7 and S2.8), which is often the case for large scale surveys (Lagier et al., 2012), or that low-abundant organisms were missed because they were below the detection limit of the technique (Pedrós-Alió, 2007). Thirdly, while sequencing depth is one aspect, PCR-related biases (e.g., GC-content) and sequencing errors (e.g., homopolymers) may also contribute to the observed differences (Harismendy et al., 2009; Berry et al., 2012; Pinto and Raskin, 2012). A GC-content deviating strongly from 50 % may induce a PCR-bias and this could explain why certain OTUs were not detected through pyrosequencing. However, calculation of the %GC of the cultivation-only sequences, in combination with the high number of such OTUs (67.8 % of the cultivation OTUs), dismissed this hypothesis in our case (Table 2.1). Although our preprocessing was done rigorously, e.g. (Huse et al., 2007; Edgar et al., 2011; Schloss et al., 2011), we cannot exclude the possibility that some erroneous sequences have slipped through (Charles K. Lee et al., 2012). Nevertheless, the limited overlap between culturing and pyrosequencing data is in line

Table 2.1. Comparison of the GC content of the cultivation-only sequences with the overall values.

	Forward sequencing (V1-V2 regions)			**Reverse sequencing (V3-V2 regions)**		
	Average (%)	**Minimum (%)**	**Maximum (%)**	**Average (%)**	**Minimum (%)**	**Maximum (%)**
All sequences	54	28	70	56	34	67
Cultivation-only[a]	52	43	63	55	51	66

[a]Sequences from cultivation that were not picked up by pyrosequencing

with observations from comparisons of cultivation and other culture-independent techniques (e.g., clone libraries) in other ecosystems (Donachie et al., 2007). High-throughput culturing (The cultural revolution, 2012) and the use of more diverse growth conditions (Delavat et al., 2012; Stewart, 2012) would probably show that the actual overlap is (much) larger than our results currently suggest. Indeed, extending the incubation time (e.g., up to three months) might reveal additional rare community members (Davis et al., 2005). Moreover, cultivation is even able to detect novel organisms where culture-independent techniques fail (Delavat et al., 2012). It has been proposed that 5000 denoised reads may be needed to describe 90 % of the alpha-diversity of 15-20.000 reads and that because of the huge bacterial diversity, almost an infinite number of individuals might need to be identified to accurately describe communities (Lundin et al., 2012).

Our comparison further confirmed that even low-abundant but widely distributed organisms can be picked up by both techniques. As an example, Figure 2.4 shows the distribution of genera in sample BB115 where, typically, the majority of genera are represented by only one or two sequences, some are moderately abundant and a few are very abundant taxa. That cultivation can pick up low-abundant bacteria may often be the result of cultivation conditions allowing enrichment of these taxa. For example, OTU 2399 (*Brevundimonas* sp., *Alphaproteobacteria*) was recovered from sample SK5 six times through pyrosequencing, while 38 times through cultivation. The ability of cultivation to pick up organisms from the

rare biosphere was also demonstrated by Shade *et al*. (2012), and these and our results show that the nutritional or cultivation requirements of these rare organisms are not necessarily extensive (The cultural revolution, 2012). In fact, *Escherichia coli* is probably the best example to demonstrate this fact. While readily cultured and even functioning as a Gram-negative model organism, it is not a very abundant organism in the human gut (Reeder and Knight, 2009).

Of the high-abundant OTUs (i.e., having more than 80 sequences) obtained by forward pyrosequencing, only three were also retrieved through cultivation (Figure S2.3). OTU2742 (*Porphyrobacter*, *Alphaproteobacteria*) was detected through pyrosequencing in samples LA3, PQ1, TM4, SO6 and WO10, and recovered by cultivation from SK5 and PQ1. A second OTU (OTU1961; unclassified alphaproteobacterium) found in BB115, BB50 and PQ1 was also found in two culture samples (SK5 and BB115). Strikingly, neither of these was found in the pyrosequence data of sample SK5. Finally, OTU2229 (*Sphingopyxis*, *Alphaproteobacteria*) was recovered from samples TM2 and WO10 through cultivation, and from samples SK5, SO6 and LA3 by pyrosequencing. In contrast, most of the OTUs frequently obtained via culturing (more than 10 sequences) were also picked up from the same sample by pyrosequencing, although generally at a lower relative abundance than through cultivation (Figure S2.1). Moreover, no OTU was shared and present at a high relative abundance in both datasets. In the reverse pyrosequencing dataset nine of the frequently recovered OTUs were also picked up by cultivation. One of these (OTU2043, unclassified alphaproteobacterium) was among the high-abundant OTUs in both techniques (Figures S2.2 and S2.5).

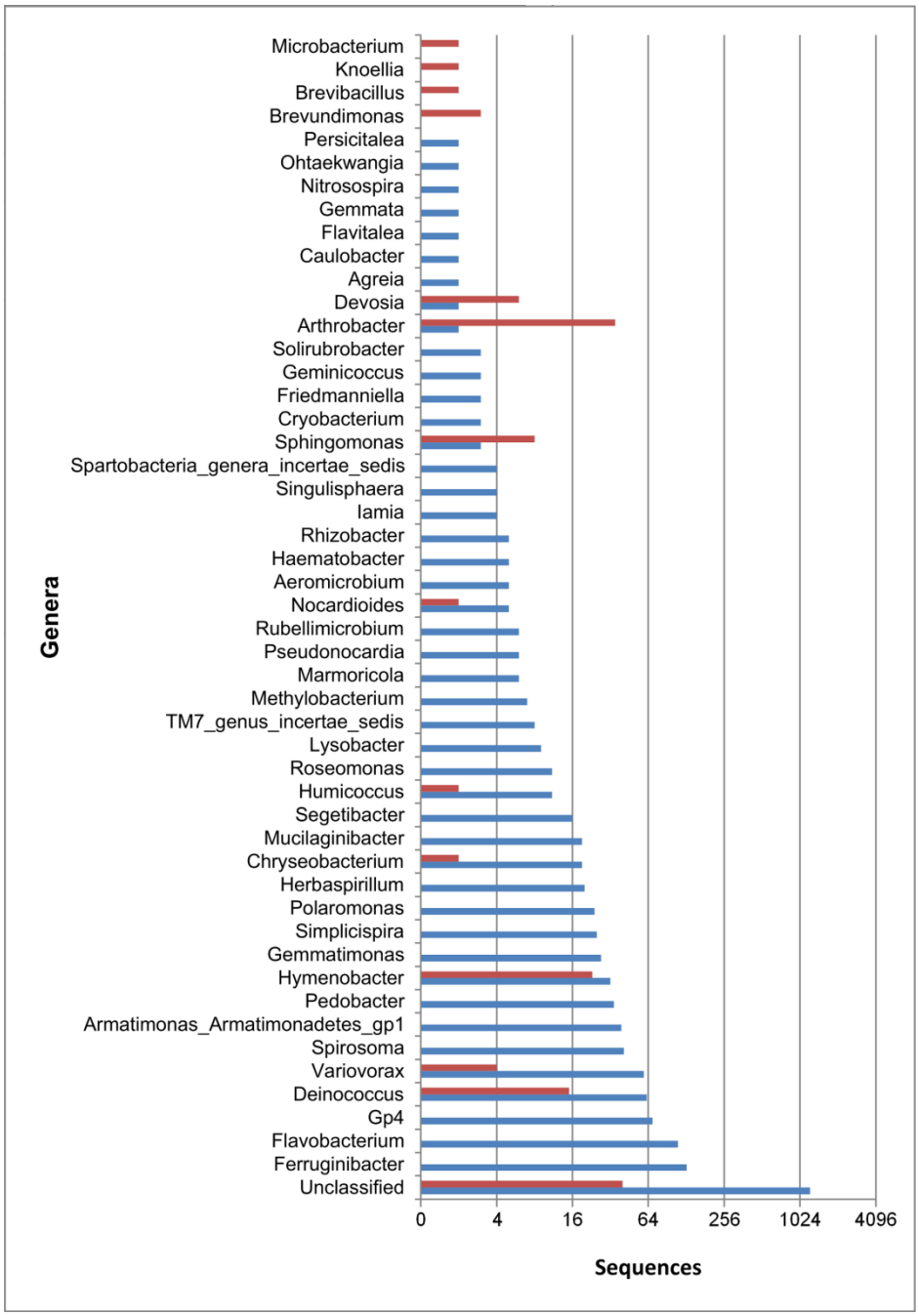

Figure 2.4. Rank-abundance plot showing the distribution of genera in a sample, illustrating the difference between techniques. Sequence numbers are plotted on a log scale. Blue bars are pyrosequencing based, red bars are cultivation based.

Notable diversity observations

While *Cyanobacteria* was the dominant phylum of photosynthetic bacteria in all samples, also the phylum *Chloroflexi* was present in all samples. Remarkably, diversity was considerably less in the forward dataset (47 OTUs including genera *Leptolinea* and *Chloroflexus*) than in

the reverse dataset (75 OTUs including *Leptolinea*, *Levilinea*, *Caldilinea*, *Heliothrix*, *Herpetosiphon*, *Dehalogenimonas*, *Sphaerobacter*). The genus *Caldilinea*, originally described for thermophilic filamentous bacteria (Sekiguchi et al., 2003; Grégoire et al., 2011), was present in all samples (Table S3.2). The phylum *Chlorobi* was much less well represented (2 OTUs in one sample).

The phylum *Planctomycetes* was also well represented in all samples: eight genera were detected, although the diversity differed between samples. Notable is the relatively frequent presence of the unusual freshwater genus *Gemmata* (Devos, 2013) with 29 OTUs found in seven of the nine samples in the forward dataset (Table S2.2).

The genus *Deinococcus* was frequently recovered in the terrestrial samples, which was also especially obvious through cultivation (BB50). Among limnetic mat samples, this genus was only recovered from PQ1 and SO6 by cultivation and was not picked up by pyrosequencing (Table S2.2). The more exposed nature of terrestrial sites may provide habitats that are particularly suited to *Deinococcus* species which are known for their resistance to radiation and desiccation (Hirsch et al., 2004).

A small number of genera were relatively frequently detected in the pyrosequencing data of both terrestrial samples (BB50 and BB115) but rarely in the seven aquatic samples: *Hymenobacter* (30 OTUs terrestrial vs. 5 OTUs in one aquatic sample), *Spirosoma* (17 OTUs terrestrial vs. 1 OTU in one aquatic sample) and *Deinococcus* (12 OTUs terrestrial samples only). Conversely, a considerable diversity of the aquatic and clinical genus *Legionella* was picked up from the aquatic mat samples (62 OTUs in the forward and 39 in the reverse dataset from 6 or 5 of the samples) while no *Legionella* was detected in the terrestrial mat samples.

Pyrosequencing allowed us to obtain a considerable number of OTUs which are as yet unidentified at the genus level (e.g., 70.27 % in the forward dataset) in addition to the

potentially new taxa already detected through cultivation (Peeters and Willems, 2011; Peeters, Ertz, et al., 2011; Peeters, Hodgson, et al., 2011; Peeters, Verleyen, et al., 2011). These might represent novel diversity adapted to the pristine and unique environment of Antarctica. This high number of novel sequences is comparable to other NGS studies in extreme and as yet understudied habitats. For example, 46 % of the sequences from an acidic Andean hot spring remained unclassified at the phylum level (Bohorquez et al., 2012). However, the high number of novel sequences might in part also be related to (i) the database used (RDP) which contains a relatively low number of sequences, but is of high quality, and (ii) the presence of artifacts that could inflate the diversity. Indeed, in view of the many possible factors that can increase the sequence diversity, pyrosequencing data are often extensively filtered to remove flawed and chimeric sequences (Huse et al., 2007, 2010; Quince et al., 2009, 2011; Edgar et al., 2011; Schloss et al., 2011). The sequence processing pipeline used here might reduce the error rate down to 0.02 % (Schloss et al., 2011). We therefore assumed that the remaining sequences are of considerable quality, and that most remaining sequencing errors would be masked by clustering. Clustering of the OTUs at 95 % similarity did not result in a large reduction of the number of OTUs or singletons (data not shown), indicating considerable diversity among the OTUs left. Our chimera filtering removed 2.5 % and 5.7 % of the total sequences in the forward and reverse data respectively. We opted not to remove the singletons and low-abundant sequences because (i) our approach already eliminated 16.6% (forward data) and 43.4% (reverse data) of the non-redundant sequences, and (ii) removing singletons may eradicate not only low quality sequences, but also biologically relevant sequences and novel taxa. In fact, 26 out of the 110 OTUs (23.6 %) shared by both pyrosequencing and cultivation were singletons in the forward pyrosequencing data that were thus readily picked up from one or more samples through cultivation. In the reverse dataset the singletons comprised 8 of the 77 overlapping OTUs (10%). These high levels indicate that

indiscriminate removal of all singletons would eliminate a considerable portion of the actual diversity.

2.6. Conclusions

Next Generation Sequencing techniques such as 454 pyrosequencing allow a much deeper sampling of microbial communities compared to the more traditional techniques. Our study revealed many unidentified OTUs and showed that the terrestrial and lacustrine bacterial diversity in Antarctica is orders of magnitude larger than previously believed. The comparison between NGS and culturing revealed that both techniques are complimentary and that only a limited number of OTUs is shared between both datasets. Although only a small number of these organisms were cultured, cultivation was able to pick up organisms from the rare biosphere, including organisms that were not recovered from pyrosequencing. With more sequencing depth and increasing read length, this may improve. It is clear that despite the ongoing technological developments, cultivation remains a useful method to uncover unknown diversity, and is currently certainly still needed for the physiological characterization and unambiguous identification of these organisms. Our comparison of forward (covering V1 and V2) and reverse sequences (covering V3 and part of V2) also revealed considerable differences in diversity obtained between variable regions and differences in the number of chimeras present. These aspects should be considered when comparing different studies.

2.7. Supplementary figures

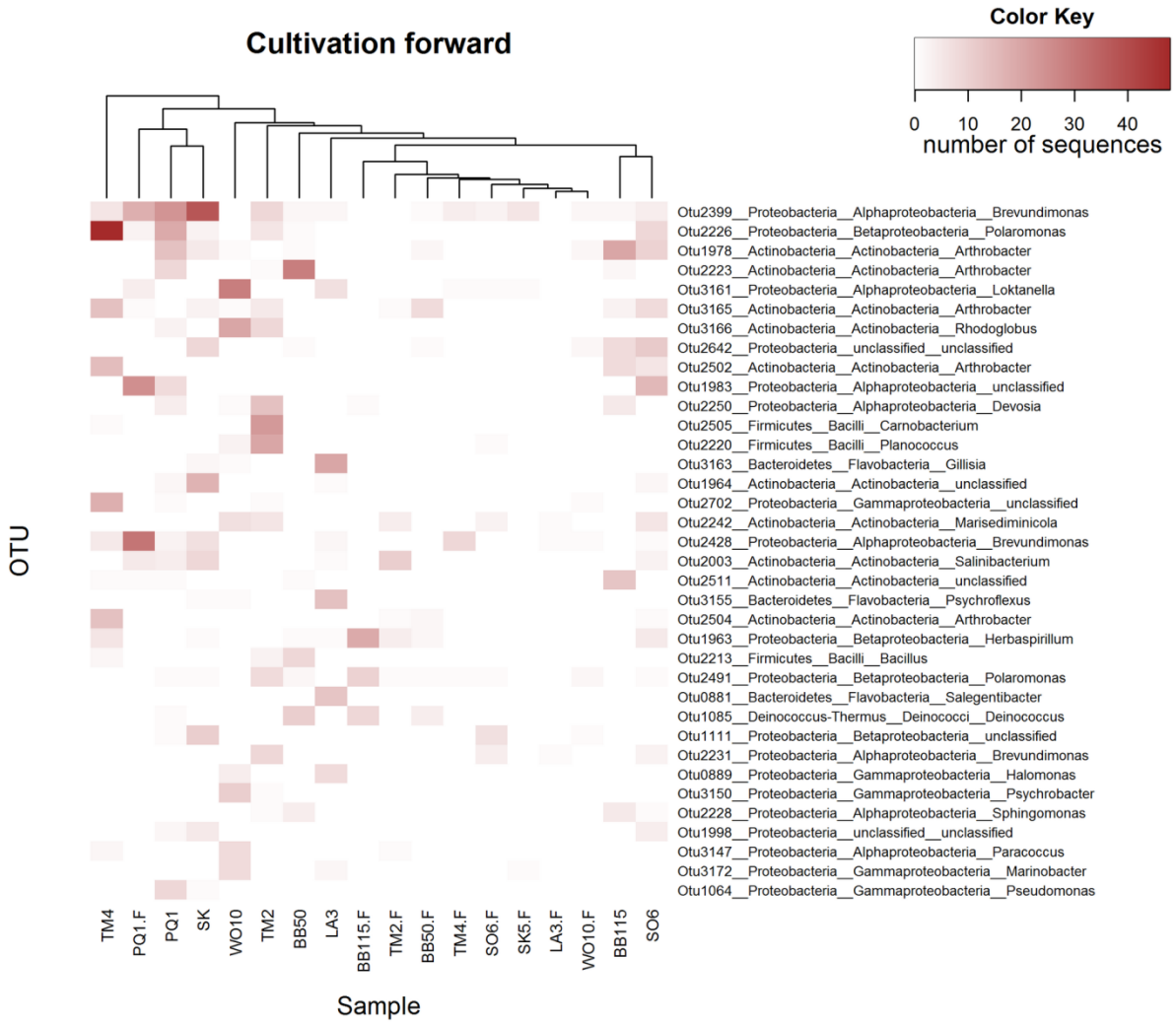

Figure S2.1. Heatmap showing the distribution of the most abundant OTUs based on the forward cultivation sequences. These high abundant OTUs are represented by at least 10 sequences. Pyrosequenced samples have the suffix .F.

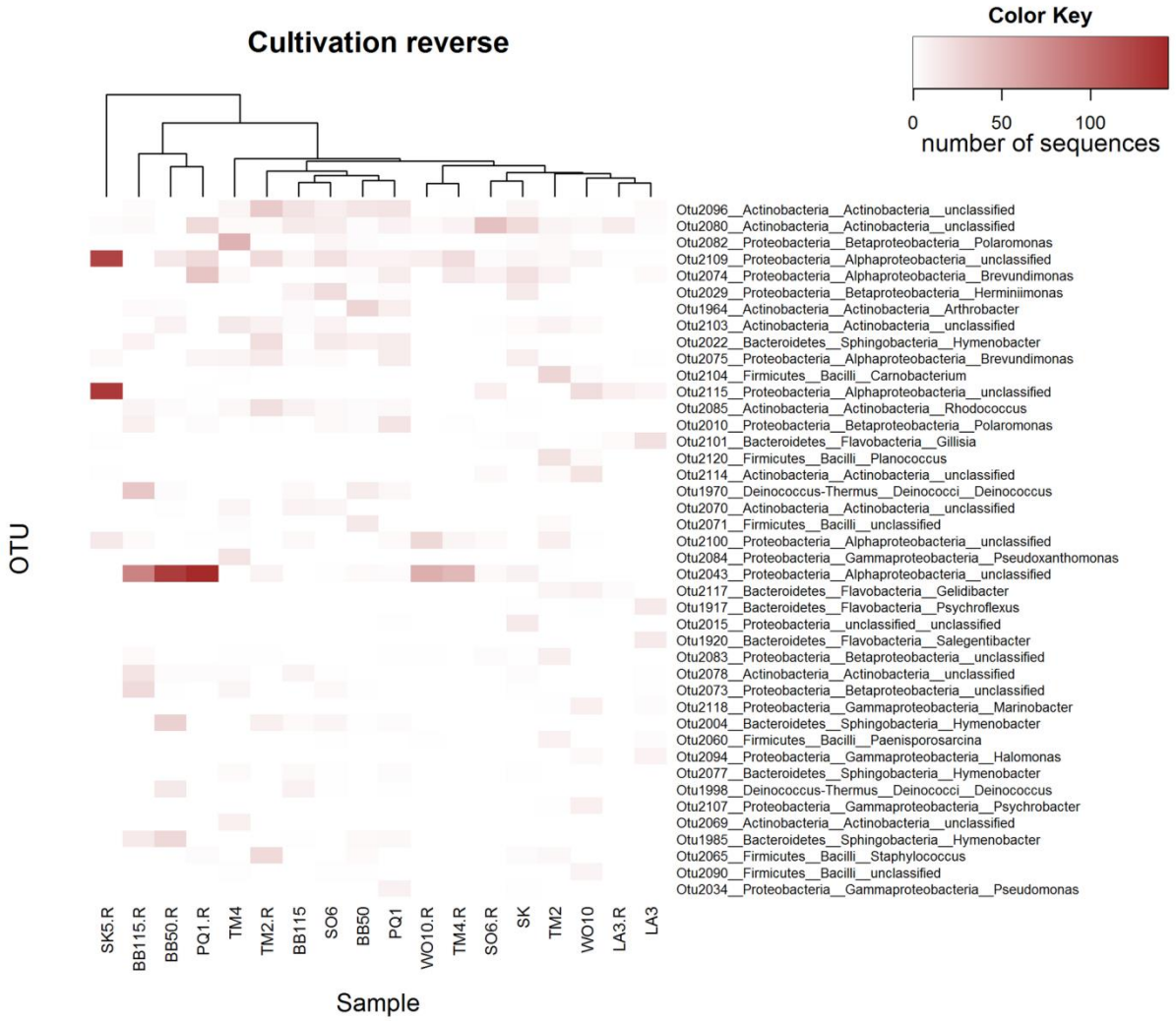

Figure S2.2. Heatmap showing the distribution of the most abundant OTUs based on the reverse cultivation sequences. These high abundant OTUs are represented by at least 10 sequences. Pyrosequenced samples have the suffix .R.

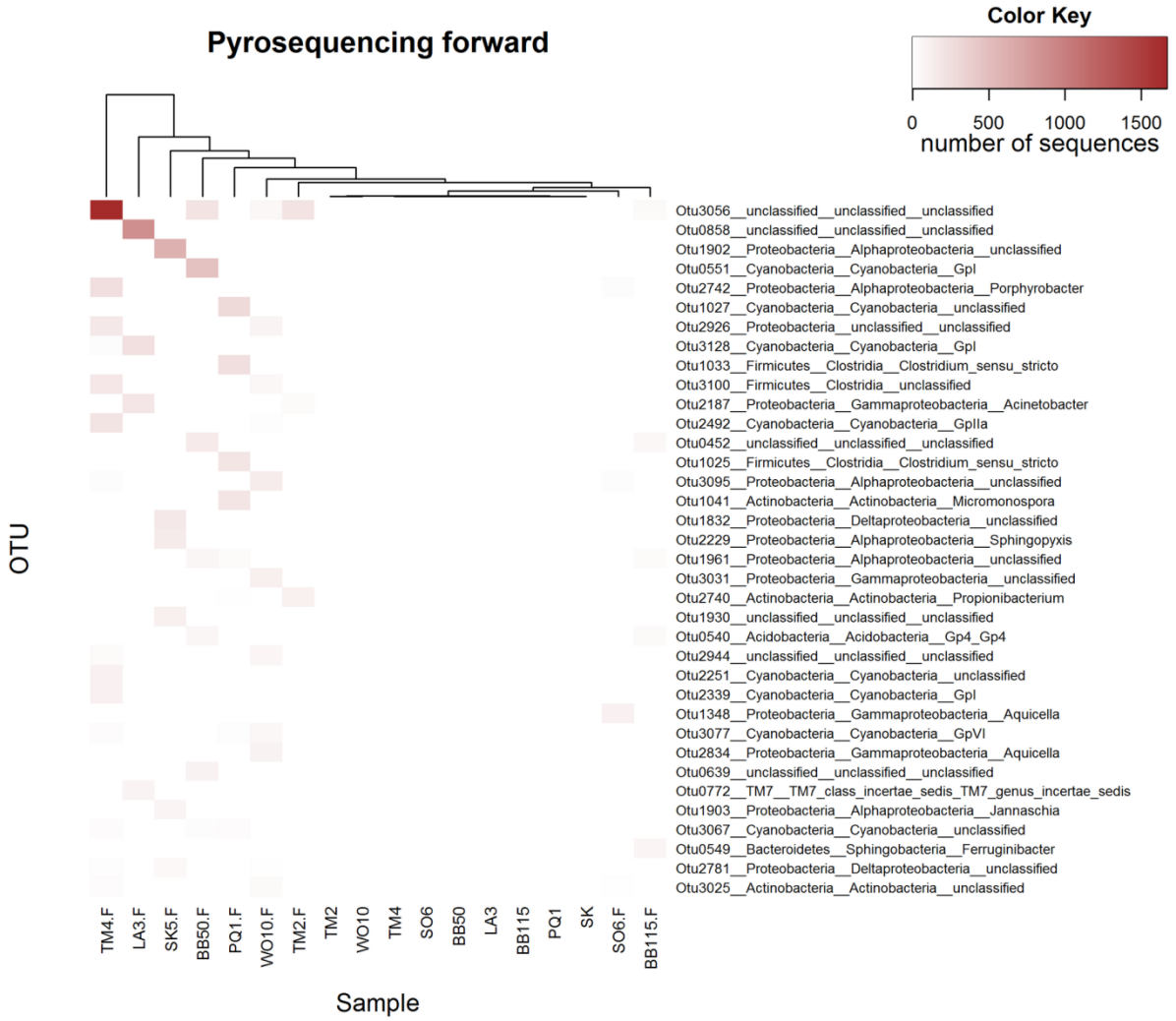

Figure S2.3. Heatmap showing the distribution of the most abundant OTUs based on forward pyrosequencing. These high abundant OTUs are represented by at least 80 sequences. Pyrosequenced samples have the suffix .F.

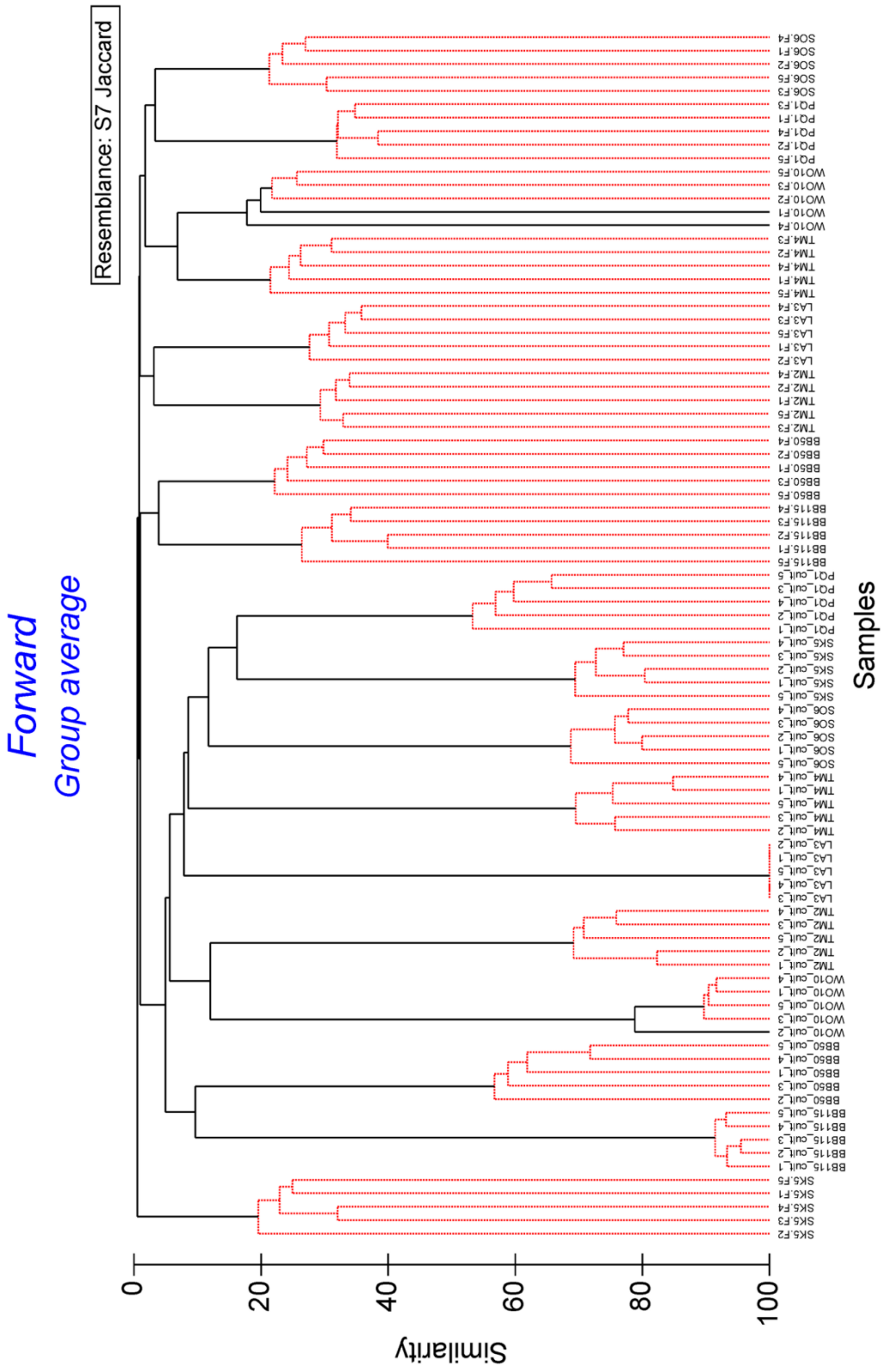

Figure S2.4. SIMPROF showing the clustering of the forward dataset. Each sample was subsampled 5 times with replacement to the lowest number of sequences (119 in cultured sample LA3). Full (black) lines are significant, dashed (red) lines are not.

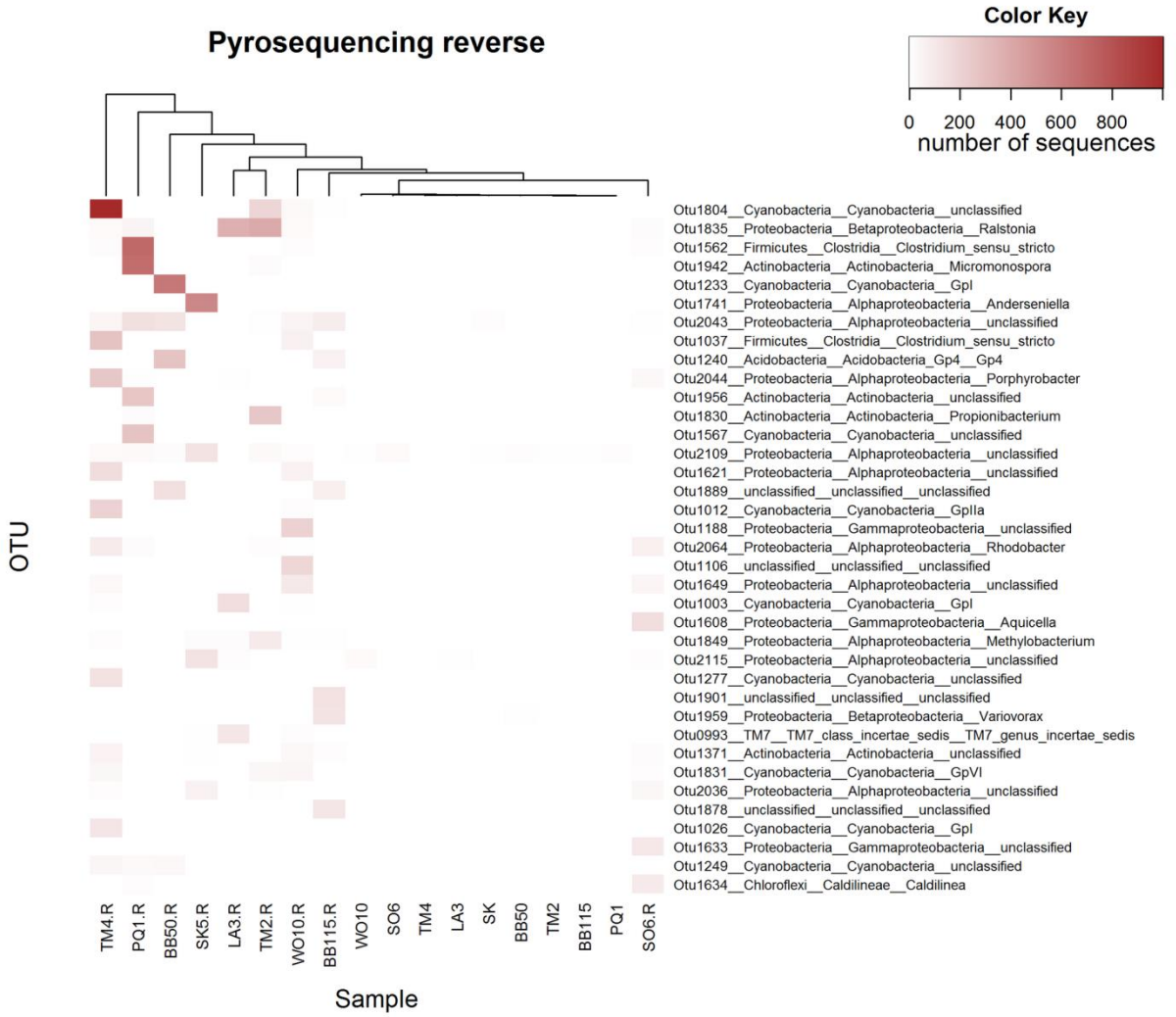

Figure S2.5. Heatmap showing the distribution of the most abundant OTUs based on reverse pyrosequencing. These high abundant OTUs are represented by at least 100 sequences. Pyrosequenced samples have the suffix .R.

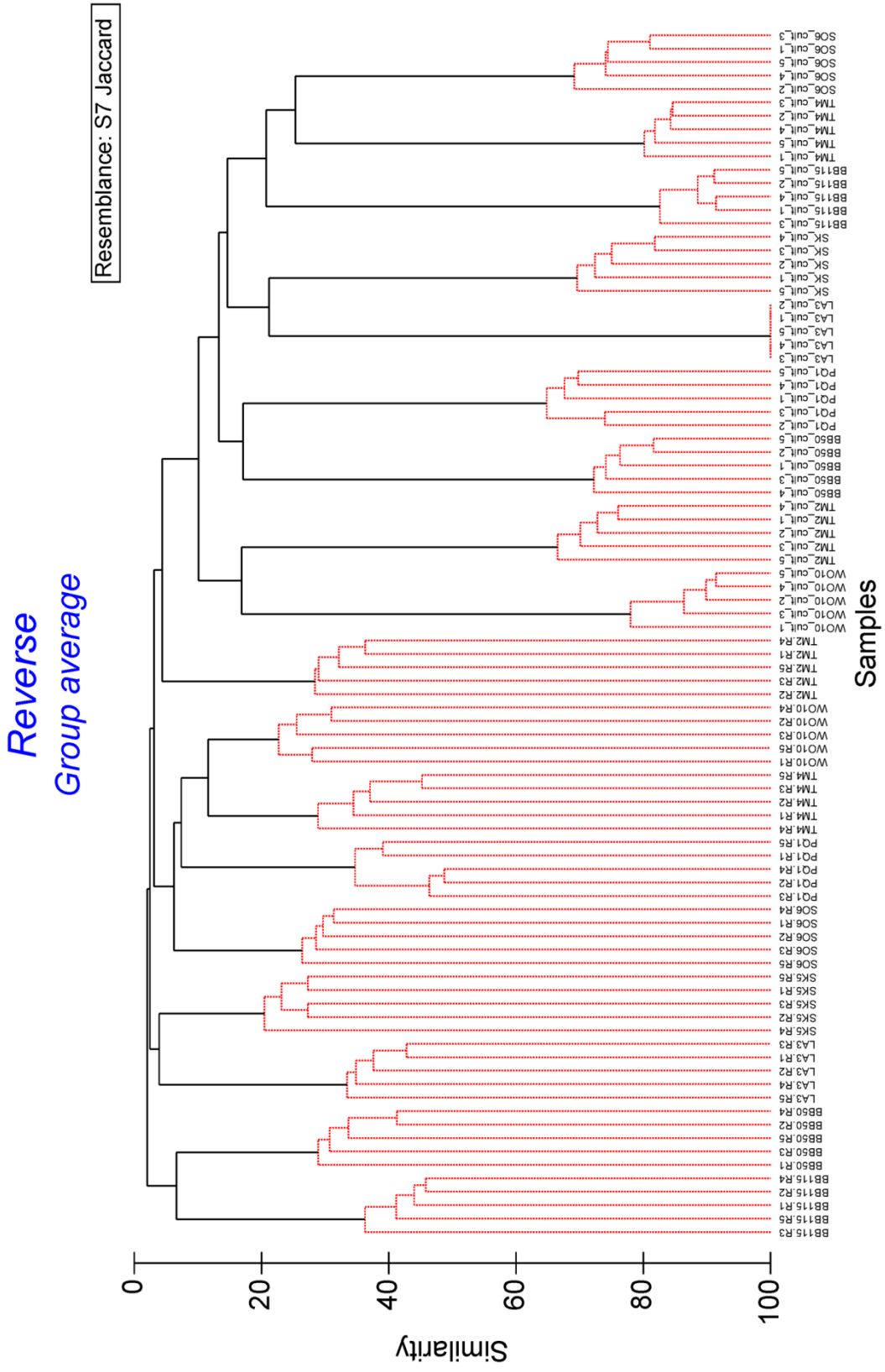

Figure S2.6. SIMPROF showing the clustering of the reverse dataset. Each sample was subsampled 5 times with replacement to the lowest number of sequences (116 in cultured sample LA3). Full (black) lines are significant, dashed (red) lines are not.

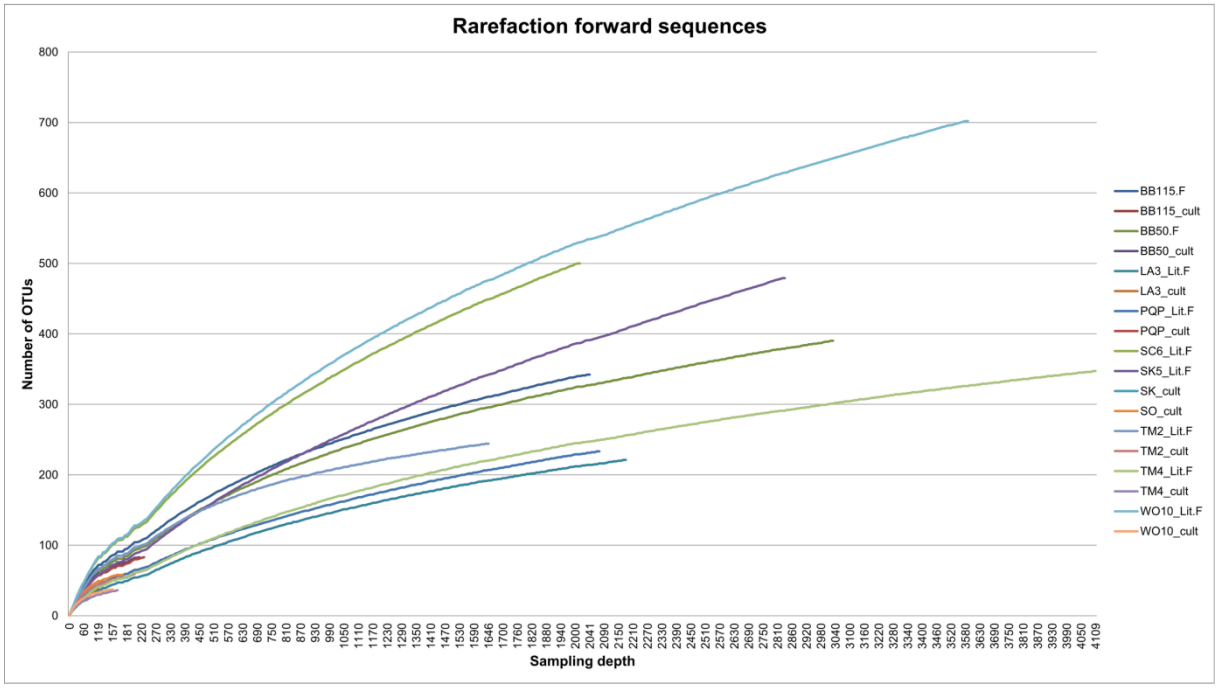

Figure S2.7. Rarefaction of the forward sequenced samples.

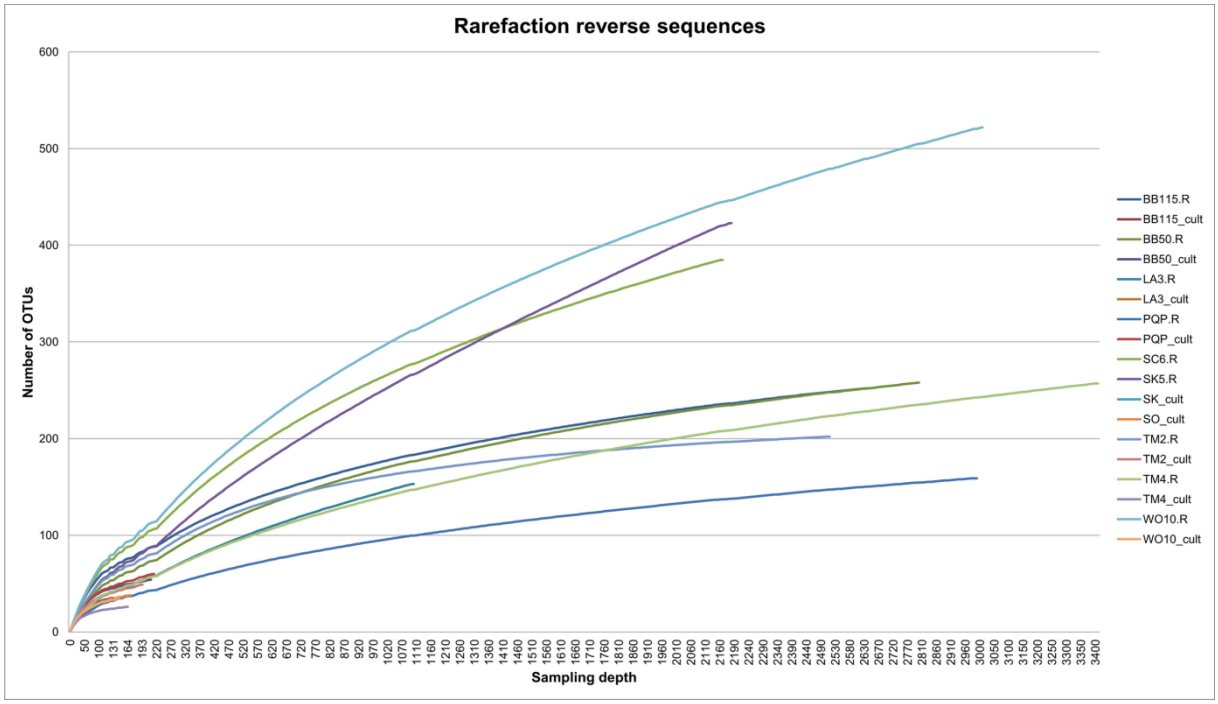

Figure S2.8. Rarefaction of the reverse sequenced samples.

Table S2.1. Comparison of the per sample chimera content (percentage) as detected using the Uchime algorithm for both pyrosequencing directions.

	Forward		Reverse	
Sample	**R**[a]	**NR**[b]	**R**	**NR**
BB115	1.4	3.2	1.3	4.9
BB50	2.6	1.8	0.8	3.0
LA3	1.7	6.7	4.9	19.3
SK5	5.8	16.6	23.0	43.3
SO6	0.5	1.0	1.8	5.8
PQ1	1.8	3.6	22.9	19.6
TM2	0.1	0.5	1.2	5.7
TM4	3.0	8.7	4.6	18.9
WO10	3.9	7.7	10.3	29.5

[a] Redundant dataset

[b] Non-redundant dataset (i.e., unique sequences)

Table S2.2. Overview of the genera recovered. The number of OTUs within each genus is shown per sample for both pyrosequencing and cultivation.

Table S2.2. is online available at:
http://journals.plos.org/plosone/article/asset?unique&id=info:doi/10.1371/journal.pone.0097564.s010

Table S2.3. Summary of the number of sequences and OTUs at the phylum level.

Table S2.3 online available at:
http://journals.plos.org/plosone/article/asset?unique&id=info:doi/10.1371/journal.pone.0097564.s011

Chapter 3. Bacterial community composition in relation to bedrock type and macrobiota in soils from the Sør Rondane Mountains, East Antarctica

Redrafted from:

Tytgat, B., Verleyen, E., Sweetlove, M., D'hondt, S., Clercx, P., Van Ranst, E., Peeters, K., Roberts, S., Namsaraev, Z., Wilmotte, A., Vyverman, W. and Willems A. (*in revision* FEMS Microbiology Ecology). Bacterial community composition in relation to bedrock type and macrobiota in soils from the Sør Rondane Mountains, East Antarctica.

Author contributions

BT, AW and EV developed the study. BT and SD performed DNA-extractions and the library preparations. BT and PC performed ARISA. BT, EV, A. Wilmotte and MS analysed the data. ZN, KP and A. Wilmotte collected the samples. BT, EV, WV, A. Willems, EVR, SR, and ZN wrote the manuscript.

3.1. Summary

The microbial biodiversity of inland Antarctic regions and the factors shaping these communities are poorly understood. We used ARISA fingerprinting in combination with the Illumina MiSeq platform to analyse the bacterial community structure in 52 soil samples from the western Sør Rondane Mountains (Dronning Maud Land, East Antarctica). The samples were taken along broad environmental gradients in an area covering nearly 1000 km², spanning differences in pH, electric conductivity, and moisture and organic matter content, two main types of bedrock (gneiss versus granite derived), and the presence of mosses or lichens. Ordination analyses revealed that total organic carbon content was the most significant variable in structuring prokaryotic communities, followed by pH, electric conductivity, bedrock type and moisture content, while spatial distance was of relatively minor importance. *Acidobacteria* (*Chloracidobacteria*) and *Actinobacteria* (*Actinomycetales*) dominated mineral soil samples situated on gneiss derived bedrock, while *Proteobacteria* (*Sphingomonadaceae*), *Cyanobacteria*, *Armatimonadetes* and candidate division *FBP* (OP11) dominated soil samples with a high total organic carbon content and were mainly situated on granite derived bedrock.

3.2. Introduction

The long-term isolation of the Antarctic continent and its history of glacial expansions and retractions since the Late Eocene resulted in truncated and strongly impoverished food webs comprising a low number of macroscopic taxa (Chown and Convey 2007; Convey 2008). Hence, terrestrial life in Antarctica is mainly microbial (Cowan et al., 2002, 2014; Vyverman et al., 2010) and largely confined to the sparse and discontinuously distributed ice-free regions (Chown and Convey, 2007). These regions are restricted to coastal lowlands, the

Transantarctic Mountains and parts of the mountain ranges that protrude through the ice sheet across the edges of the continent (i.e., nunataks) (Pugh and Convey, 2008), and are believed to have acted as refugia for terrestrial biota during Neogene and Pleistocene ice ages (Convey 2008; Pugh and Convey 2008). While it was long believed that soils in the most inland, high altitude regions were devoid of life, it is now well-accepted that even the most arid soils with virtually no organic matter are inhabited by microbial communities (Magalhães et al., 2012; Tytgat et al., 2014). Despite their ecological importance (Wynn-Williams, 1996; Brinkmann et al., 2007), the microbial biodiversity in these ice-free regions and the factors structuring their (meta-)communities remain poorly studied.

The most intensively studied terrestrial ecosystems in Antarctica are those in the McMurdo Dry Valleys (MDV) in South Victoria Land (see Cary *et al.* 2010 for a review), while the microbial diversity in other and particularly more inland locations remains poorly known. However, findings obtained in one region cannot be simply extrapolated to others, because, for example, regional differences in environmental gradients and (micro-)climatic conditions. The polar desert soils in the MDV differ in at least one aspect from those in more inland nunataks, namely their carbon source. More in particular, while in high altitude regions *Cyanobacteria*, algae, mosses and lichens, and probably also autotrophic bacteria are responsible for the majority of the carbon fixation (Brinkmann et al., 2007; Namsaraev et al., 2010; Fernández-Carazo et al., 2012; Magalhães et al., 2012; Ertz et al., 2014; Tsujimoto et al., 2014; Tahon et al., 2016b), the carbon sources in the soils of the MDV are of multiple origins (Cary *et al.* 2010). Apart from recent biological C-fixation, organic matter in some soils of the Dry Valleys is in part derived from old lacustrine sediments which were deposited in large lakes that existed during the Last Glacial Maxima (Hall et al., 2010). In other parts of the region, the carbon stock is influenced by recently deposited cyanobacterial mats derived from the littoral zone of the present-day lakes and rivers or melt-water channels, or by marine

inputs of organic matter (e.g., seal carcasses and sea spray) in the more coastal regions (Cary *et al.*, 2010). By contrast, with a few exceptions (Hodgson et al., 2010), lakes are absent in high altitude nunataks, and the soils were never part of a lacustrine (lake or river) environment. Moreover the marine influence is restricted to a small number of marine birds nesting along mountain slopes up to 200 km inland (e.g., Snow Petrel *Pagodroma nivea*; Ohyama and Hiruta, 1995; Peeters 2011). It follows that the environmental gradients shaping the microbial communities in these inland nunataks might be different to those in the MDV.

At least 14 phyla of bacteria have been recovered from MDV soils, with *Proteobacteria*, *Actinobacteria*, *Acidobacteria*, *Gemmatimonadetes* and *Bacteroidetes* being amongst the most dominant ones (Cary 2010). Recently, Kim (2015) reported similar observations from Terra Nova Bay, North Victoria Land, while *Chloroflexi* appeared to dominate in the more acidic soils. This agrees with previous findings that pH is the most important factor in structuring soil bacterial communities at lower latitudes (Fierer and Jackson, 2006). However, because of the low buffering capacities of Antarctic (mineral) soils (Cannone and Guglielmin, 2009), and the close association of microorganisms with their environment (Vrionis et al., 2013), it can be hypothesised that the soil geochemical characteristics also enforce a high selective pressure on the bacterial communities in these soils, and that different bedrock types support different bacterial communities (Barton et al., 2007). In addition to pH, it appears that electric conductivity (EC) and the organic carbon content or C/N-ratio are the main environmental factors structuring terrestrial communities, while moisture availability has varying effects (Chong et al., 2012; Magalhães et al., 2012; Geyer et al., 2014; Kim et al., 2015). Furthermore, the scattered presence of plants provides highly localized nutrient sources and protection for microbial communities in the otherwise oligotrophic and harsh Antarctic conditions. For example, it has been shown that plant cover influenced the ground thermal regime (Cannone et al., 2006; Cannone and Guglielmin, 2009) and resulted in a higher bacterial richness

compared to bare mineral soils along a latitudinal gradient in the Antarctic (Yergeau *et al.* 2007b).

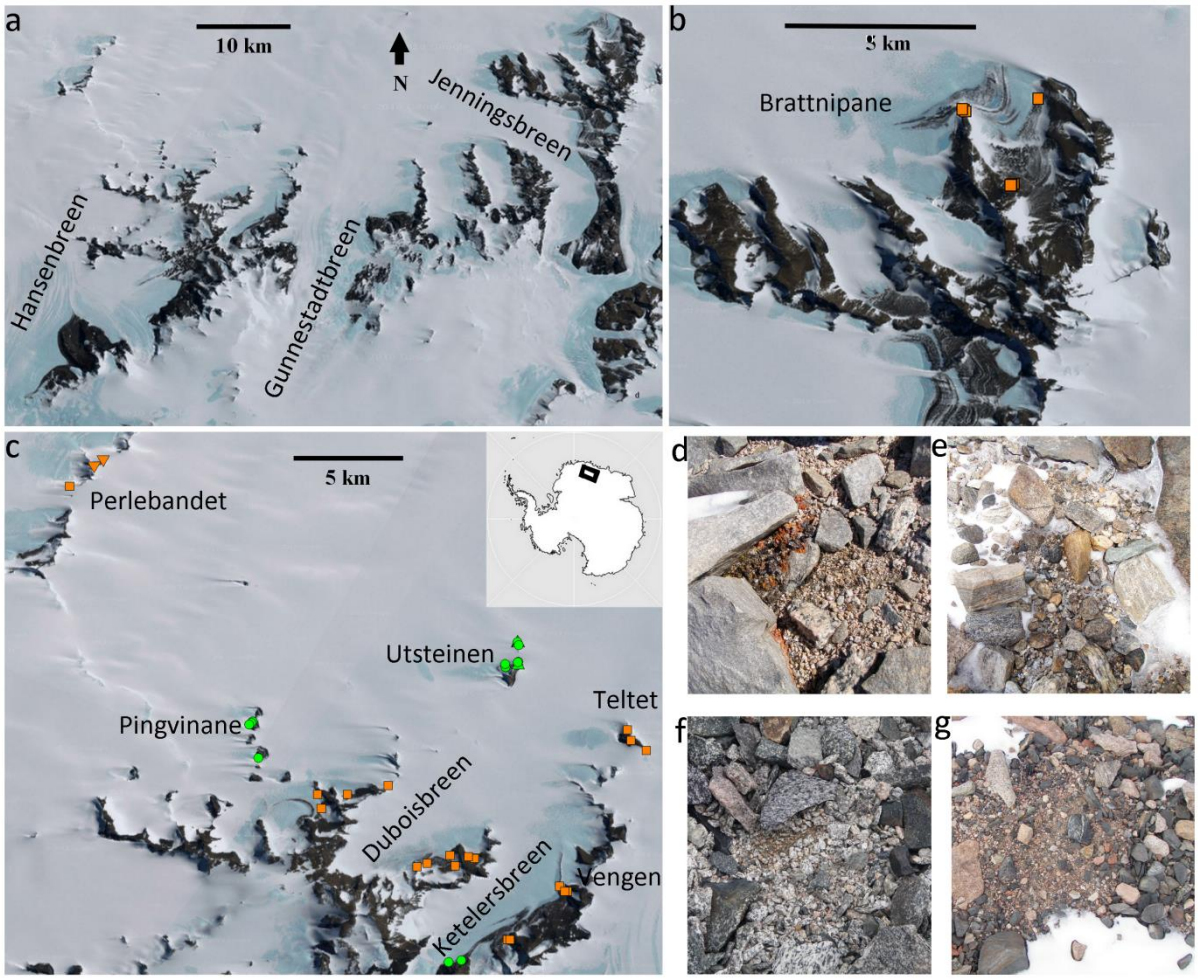

Figure 3.1. a. Map of the western Sør Rondane Mountains within Dronning Maud Land, East Antarctica, with the major outlet glaciers. b. detail of the Brattnipane region (BP) with the sampled locations. c. detail of other sampled regions included the three nunatak sites Utsteinen (UN), Teltet (TE) and Pingvinane (PA), and four areas within the main mountain range, namely the east (DE) and west (DW) sides of the Duboisbreen valley, a granite outcrop south of Ketelersbreen (KB) and Vengen (VE). Granite bedrock is indicated by the green symbols, the gneiss derived samples by the orange symbols. Macrobiota containing samples are triangles. Images d to g are examples of different bedrock at different locations. d. sample UT129 with the presence of orange lichens. e. BP156. f. KB112. g. VE137.

Here we aimed to study the effects of bedrock type, the presence of vegetation, the concentration of inorganic and organic carbon, the electric conductivity, the moisture content, and the pH on the bacterial community structure and diversity in inland Antarctic soils. To

achieve this, we selected 52 soil samples covering a broad range of environmental gradients from nine nunatak regions situated on either gneiss or granite derived bedrock in the western part of the Sør Rondane Mountains (SRM; Fig. 3.1) (Dronning Maud Land) and used Illumina sequencing and Automated Ribosomal Intergenic Spacer Analysis (ARISA) fingerprinting (Fisher and Triplett, 1999) to assess differences in their bacterial community structure.

3.3. Materials and methods

The SRM (22°E-28°E, 71°30'S-72°40'S) are a 220 km long, east-west oriented mountain range, situated about 200 km inland (Pattyn et al., 2010; Osanai et al., 2013), with peaks reaching 3300 m a.s.l. (Gorodetskaya et al., 2013). The bedrock is mainly metamorphic (e.g., gneiss) with some plutonic outcrops (granite) (e.g., Pingvinane and Utsteinen, Fig. 1). Near the Utsteinen nunatak, in the escarpment zone north of the SRM, the Belgian Princess Elisabeth Station was built, where meteorological data has been collected since 2009 (Pattyn et al., 2010; Gorodetskaya et al., 2013). Because of the orientation of the mountains, the ice flow from the polar plateau is largely blocked and forced around the range. Katabatic winds are mainly confined to the major outlet glacier to the west (Hansenbreen) or those cutting through the eastern part of the western SRM (Gunnestadtbreen, closest to Utsteinen, and Jenningsbreen, west of the Brattnipane region) (Fig. 1a), leaving the area directly north of the western SRM rather sheltered. As a result, with increasing distance from the plateau more favourable conditions for macrobiotic lifeforms (mosses and lichens) occur. The SRM experience relatively mild winters (so called coreless winters with hardly any variation in temperature during nearly five months (Bargagli, 2005), typical for the Antarctic inland) with average winter temperatures of -23 °C at Utsteinen, while summer averages are about -8 °C (Pattyn et al., 2010). However, large fluctuations in daily air (4 to 5 °C) and surface (11 to

14 °C) temperatures can occur (Gorodetskaya et al., 2013). Precipitation for 2012 was reported to be 110 ± 20 mm water equivalent year^{-1} (Gorodetskaya et al., 2015) and net snow accumulation at Utsteinen is highly variable between years (Gorodetskaya et al., 2013, 2015).

During sampling campaigns in the austral summers of 2009 and 2010, 52 terrestrial samples consisting of c. 10-40 g of the upper 3 cm of soil were collected from nine ice-free areas in the western Sør Rondane Mountains, with distances between samples ranging from less than 1 m up to a maximum of ~65 km. Samples were selected according to bedrock type (gneiss or syenite-granite derived), macrobiotic content (presence or absence of moss, lichen, algae and/or arthropods) and geographic location (Fig. 3.1). Thirty-one gneiss samples were collected from the Perlebandet (PB) nunataks in the northwestern most region of the SRM; the east (DE) and west (DW) sides of the Duboisbreen valley; the Vengen (VE) subrange of the Vikinghøgda; the Teltet nunatak (TE) north of Vengen; and the more distant Brattnipane region to the east. In total, 21 granite derived samples were collected from three regions (two nunatak sites outside the main mountain range (Pingvinane (PA) and Utsteinen (UT)), and a granite outcrop (KB) to the south of Ketelersbreen and Vengen). Mosses and lichens were only recovered from the UT and PB regions, while arthropods were also observed in KB, DE and PA.

All samples were collected aseptically using a spatula, immediately frozen and transported and stored frozen at -20 °C until further processing. Coordinates and elevation were recorded using a handheld GPS. Most samples consisted of only gravel and finer weathered rock material (further referred to as abiotic or mineral samples), while 10 samples contained lichens or moss, or were taken directly below lichen or moss biomass. In 10 samples, a darker, soil-like matter (i.e., an accumulation of organic matter) was found, but only six of these coincided with the presence of lichen or moss. Collembolans or mites were observed in five samples, of which four were mineral soils. Macroscopic algae were observed in one sample.

In this study, only the samples in which lichens, mosses or algae were observed will be referred to as macrobiotic samples.

Prior to analyses, the soil samples were homogenized by shaking the tubes containing the defrosted sample. Total carbon (TC) was determined by combustion of a subsample, while inorganic carbon (IC) was determined through acidification of a second subsample and subsequent combustion at a lower temperature. In both instances, CO_2 was subsequently measured through infrared spectrometry using a Shimadzu TOC5050A analyser with an SSM5000A module. Total organic carbon (TOC) was calculated by subtracting IC from TC. Prior to pH and electric conductivity measurements, homogenized soil subsamples were dried at 40 °C for at least 24 h until no further mass reduction was observed, after which the moisture content was calculated. To measure pH and electric conductivity, 15 ml of distilled water was added to 3 g of the sieved soil fraction of < 2 mm (ISO 10390 and ISO 11265). Electric conductivity was measured using an LF96 Microprocessor Conductivity Meter (WTW, Germany) at 20 °C, and pH was measured after electric conductivity on the same preparation with a P902 Multi-channel Analyser (Consort, Belgium). For each sample, the pH and EC were measured in two different sub-samples and the average of both measurements was used in subsequent statistical analyses.

DNA extraction and sequencing

Of each homogenized sample, 1 g was used for DNA extraction. First, extracellular DNA was removed following Corinaldesi, Danovaro and Anno (2005) and actual DNA extraction was performed according to Zwart (1998). PCRs targeted the V1–V3 hypervariable regions (~528 bp) and were performed in duplicate. Primers pA (AGAGTTTGATCCTGGCTCAG, positions 8–27) (Edwards et al., 1989) and BKL1 (GTATTACCGCGGCTGCTGGCA,

positions 536–516) (Cleenwerck et al., 2007) were modified with adapters to complement the Nextera XT index kit (Illumina, USA), and HPLC purified (IDT, Belgium). Each reaction consisted of 2.5 µl 10x buffer (High Fidelity PCR system, Roche, Switzerland), 2.5 µl dNTPs (10 mM) (Life Technologies, Belgium), 0.5 µl of each primer, 0.1 unit of High Fidelity Hot Start polymerase (Roche, Switzerland), 0.5 to 2 µl of DNA template, and was adjusted to a final volume of 25 µl using sterile HPLC water. An initial denaturation step of 3 min at 95 °C was followed by 27 cycles of 30 s at 95 °C, 45 s at 55 °C, and an extended elongation of 3 min at 72 °C in order to reduce the number of chimeric sequences (Engelbrektson et al., 2010). Lastly, a final elongation of 10 min at 72 °C was executed. PCR products were pooled and subsequently purified using a slightly modified Ampure beads XT (Agencourt, Beckman Coulter, USA) protocol, where only 0.8 volumes of beads were used and DNA was resuspended in milliQ water. Tagging was subsequently done with the Nextera XT indices in an eight-cycle version of the amplicon PCR with the indices replacing the primers. PCR products were purified as described above, with DNA resuspended in a 0.10 M 8.5 pH Tris buffer. DNA integrity was checked using a BioAnalyzer (Agilent Technologies, USA) and concentrations were measured using the High Sensitivity array for Qubit (Thermo Fisher Scientific, USA), after which the samples were pooled equimolarly. Amplicons were sequenced on an Illumina MiSeq platform giving 2x300 bp paired end reads. The run also included a blank (a mixture of several PCRs' negative controls), two mock communities (see Addendum II) to check overall run quality and benchmark processing variables, and two replicate samples. The replicate samples included one sample which is analysed in every Illumina run done for the both UGent labs to check for consistency between different analysis, and one sample that was run twice. Sequences were deposited to the NCBI Sequence Read Archive under accession number SRP067111.

ARISA

One out of every five DNA extractions was performed in duplicate to assess the robustness of the DNA extraction method and ARISA. Internal Transcribed Spacer 1 (ITS1) amplification was performed using the universal primers 16S/p2 (CTTGTACACACCGCCCGTC, positions 1390–1407) and 23S/p7 (GGTACTTAGATGTTTCAGTTC, positions 188–208) (Gürtler and Stanisich, 1996), the former with a 6-FAM (carboxyfluorescein) label (Sigma-Aldrich, USA). The PCR mixture consisted of 5 µl 10x buffer (Life Technologies, Belgium), 5 µl (2 mM) dNTP (Life Technologies, Belgium), 1 µl (25 mM) $MgCl_2$ (Qiagen, Germany), 0.8 µl (20 µg $µl^{-1}$) BSA (Roche, Switzerland), 2.5 µl (10 µM) of each primer, 1.25 µl (1 u $µl^{-1}$) Taq polymerase (Life Technologies, Belgium), 2 µl template DNA and 29.95 µl sterile milliQ water. An initial step of 5 min at 94 °C was followed by 35 cycles of 1 min denaturation at 94 °C, 1 min of annealing at 56 °C and an elongation step at 72 °C for 1 min. After a final elongation of 5 min at 72 °C, the amplified products were purified using filter plates (NucleoFast 96 PCR, Machery-Nagel, Germany) and a Tecan Workstation 200 (Switzerland), and the concentration was adjusted to 25 ng $µl^{-1}$ using a NanoDrop 2000 (Thermo Fisher Scientific, USA). Next, a mixture of 0.2 µl GeneScan 1200 LIZ Size Standard (Applied Biosystems, USA), 7.8 µl (1 u $µl^{-1}$) Hi-Di formamide (Life Technologies, Belgium) and 2 µl of DNA template was made, centrifuged, heated for 4 min at 95 °C and subsequently cooled on ice for 2 min. ITS1 fragments were separated on an ABI Prism 3130xl Genetic Analyzer (Applied Biosystems, USA) for 108 minutes at 8.5 kV.

Data processing

Forward and reverse reads of the Illumina sequencing were paired using the *fastq_mergepairs* command in USEARCH (Edgar, 2010). We used the mock communities to determine the variables giving the best approximation to the original input set of sequences. The minimum

length was set to 450 bp, and the maximum length to 570 bp. No ambiguous bases (N) were permitted, and a mismatch of at most 6 bp was allowed for the overlap of paired reads. Further sequence quality filtering and processing was done using USEARCH, with a minimal Phred score set to Q20 and a maximum expected error of 0.5. OTU clustering was done with UPARSE (Edgar, 2013) and *de novo* chimera filtering with UCHIME (Edgar et al., 2011). An OTU cutoff of 97 % was used to cluster sequences, and the May 2013 GreenGenes training set (DeSantis, Hugenholtz, Larsen, et al., 2006; Schloss, 2010; McDonald et al., 2012) was used with the naïve Bayesian classifier (Wang et al., 2007) implemented in Mothur (Schloss et al., 2009) to infer a taxonomy for these OTUs. Non-bacterial, chloroplast and mitochondrial OTUs were removed prior to downstream analysis.

ARISA fragment lengths (AFL) were analyzed with GeneMarker v2.2.0 (SoftGenetics, USA). Community fingerprints were checked for quality and were removed when the size standard showed errors or when the ITS1 peaks showed anomalies (e.g., negative values or no peaks). After manually checking baseline noise levels, we set a minimal intensity threshold of 50 RFU (Relative Fluorescence Units). Next, peaks were binned to deal with size-call uncertainty (Brown et al., 2005), i.e. 2–4 bp bins for AFLs up to 700 bp, 5 bp bins from 700 to 1000 bp, and bins of 10 bp for larger fragments lengths. Binning is especially necessary for larger fragments (Popa et al., 2009). We set a minimal AFL of 350 bp following the observation of Kovacs, Yacoby and Gophna (2010) on minimum fragment lengths in bacterial genomes and in line with the fact that in the vast majority of our AFLs were minimally between 459 and 500 bp long. By doing so, we also eliminated the noise observed in some samples associated with smaller sized fragments. We did not eliminate shoulder peaks explicitly, as they were generally contained within a bin.

Statistical analyses

The Pearson correlation between abiotic variables was tested. TC showed a high correlation with TOC (Pearson's r = 0.99, P = 0) and was removed from the analysis. Values of the remaining abiotic variables were checked for normality (Shapiro-Wilks test) and homoscedasticity (Bartlett's test). Carbon and moisture content were logit transformed, while electric conductivity was log transformed. Two-sided t-tests were performed to detect significant differences in the values of abiotic variables and either bedrock or the presence of macrobiota. For pH, a Welch t-test was used because this is robust for non-normal data. A factorial ANOVA was used to investigate the potential interaction between bedrock type and macrobiota content (Table 3.1). In addition, an ANOVA with a Tukey *post hoc* test (Table S2) was used to investigate differences in abiotic factors between sample classes situated on different substrates (i.e., mineral gneiss (n = 29) or granite (n = 13) derived samples and macrobiota containing gneiss (n = 2) or granite (n = 8) derived samples).

To investigate the effects of the abiotic factors on the observed richness, the number of OTUs in the NGS data was recalculated based on the subsampling to the lowest number of reads (1123) in the dataset. The Pearson correlation between abiotic variables or macrobiota and the Shannon and inverse Simpson diversity indices, and the Chao1estimator were calculated.

Ordination and cluster analyses were used to assess differences in community structure between the samples and to identify those factors explaining the biotic turnover patterns between the samples. For the multivariate analysis of the ARISA data, one sample was selected randomly out of the replicate samples. The reproducibility of this approach as well as the robustness of the DNA extraction method was confirmed by calculating Pearson product moment correlation coefficients between ARISA profiles of replicate samples in Bionumerics (Applied Maths, Belgium). The multivariate analyses involved two approaches. First, a

Table 3.1. Factorial ANOVA showing the effect of bedrock type, the presence of macrobiota or their interaction on environmental parameters, OTU richness or the relative abundances of the most occurring phyla. Only significant (bold) or nearly significant ($P < 0.1$) effects are displayed.

	bedrock		macrobiota		interaction	
	F	*P*	*F*	*P*	*F*	*P*
pH	**18.1**	**< 0.001**	3.8	< 0.1	2.8	< 0.1
Electric conductivity			**4.87**	**< 0.05**		
TOC	**27.47**	**< 0.001**	**34.65**	**< 0.001**		
Chao1					3.05	< 0.1
Moisture content			**18.38**	**< 0.001**		
Actinobacteria	**7.2**	**< 0.01**	**5.64**	**< 0.05**		
Armatimonadetes	**6.25**	**< 0.05**				
Bacteroidetes			**4.44**	**< 0.05**	**5.94**	**< 0.05**
FBP	**11.77**	**< 0.01**	**7.71**	**< 0.01**		
Gemmatimonadetes			2.88	< 0.1		
Planctomycetes			**10.68**	**< 0.01**	**6.72**	**< 0.05**
Proteobacteria	**6.39**	**< 0.05**	**10.82**	**< 0.05**		
Verrucomicrobia			**4.1**	**< 0.05**		

TOC: Total organic carbon

SIMPROF analysis (Clarke et al., 2008) with 999 iterations, alpha set to 0.05 and average clustering using the Bray-Curtis distance, was used for clustering the samples based on their bacterial community structure. SIMPROF is a permutation-based procedure that ranks the pairwise similarities in each group and tests the null hypothesis that samples were all drawn from the same species assemblage. Second, redundancy analyses (RDA) were used to identify the environmental and spatial factors explaining variation in community structure based on a variation partitioning approach. Apart from the biotic datasets (ARISA and NGS based), this required two other datasets, namely the one with the environmental variables described above and one with spatial factors. First, the biotic data were Hellinger transformed because this has been shown to be a valid data transformation when analysing variation between communities at individual sites using ordination analyses (Legendre et al., 2005; Peres-Neto et al., 2006). The matrix with the spatial variables contained Principal Coordinates of Neighbour Matrices

(PCNMs) of the geographic coordinates of the samples (Borcard and Legendre, 2002; Dray et al., 2006). These are positive eigenvalues obtained by principal coordinate analysis (PCoA) of a truncated matrix of Euclidian distances between the sampling sites. Partial RDAs were subsequently run using the forward selection procedure following the protocols described in Peres-Neto *et al*. (2006) in order to identify the factors significantly explaining the variation in bacterial community structure, and to calculate the amount of variation uniquely explained by the environmental factors, the spatial factors, and the overlap between both sets of predictors. All downstream analyses and statistics for both datasets were performed in R (R Core Team, 2015), with the packages *Vegan* 2.3-0 (Oksanen et al., 2015) and *clustsig* 1.1 (Whitaker and Christman, 2014) for the redundancy analysis (Hellinger transformed data and scaled abiotic variables), the calculation of the PCNM variables (Borcard et al., 2004) and the diversity metrices

3.4. Results

SOIL PROPERTIES

A higher abundance of macrobiota was observed in the nunatak sites most distant (i.e., 10-20 km north) from the plateau with mainly granite derived soils. The granite derived samples also had a significantly higher average TOC content (t-test; $P < 0.001$) than the gneiss derived samples (2.27 ± 6.48 % vs. 0.17 ± 0.35 %) (see also Table 3.1). Sample UT50 had an extremely high TOC content (30.6 %), which could be attributed to the fact that it consisted nearly completely of moss. However, the difference in TOC content between both bedrock types remained significant after removing this outlier (t-test; $P < 0.001$). The TOC content was also higher in the samples containing macrobiotia (4.72 ± 9.26 %), compared to mineral soils (0.13 ± 0.18 %; t-test; $P < 0.005$). These patterns remained significant when further

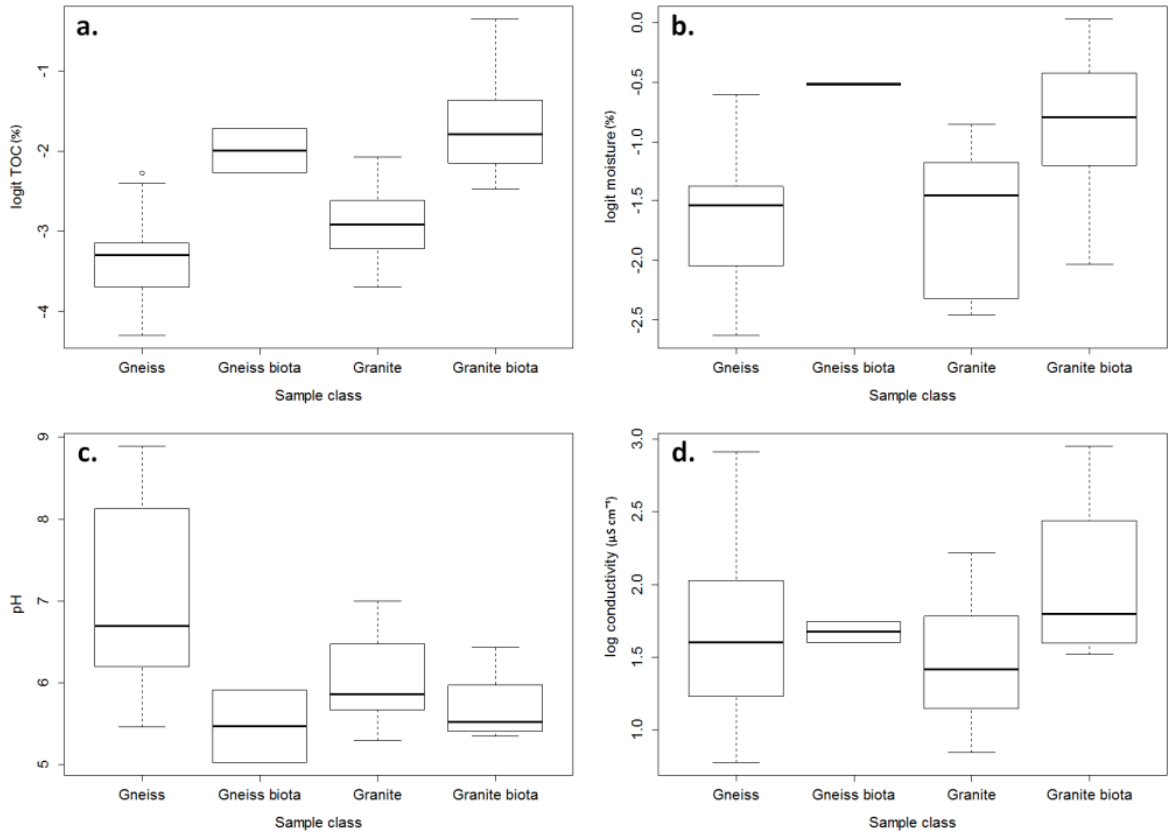

Figure. 3.2. Comparison of abiotic variables between the different sample classes. TOC, pH and moisture content differed significantly ($P < 0.001$) between, but not within, the biotic-abiotic classes (ANOVA, Tukey post-hoc test). Electric conductivity showed no significant difference between classes.

splitting up the samples in classes according to bedrock and mineral or macrobiota content (Fig. 3.2a; Table S3.1). The moisture content showed large differences between samples, ranging from 0.23 % to 51.6 %, and was on average higher for granite derived samples (9.69 ± 12.29 % vs. 4.77 ± 6.15 %), but this difference appeared to be insignificant. The difference in moisture content between mineral samples and those containing macrobiota, however, was highly significant (t-test; $P < 0.001$), with the extreme value of 51.6 % again being attributable to sample UT50. On average, samples in which macrobiota were present contained 19.6 ± 14.8 % moisture, compared to 3.7 ± 3.89 % for the mineral soils. These patterns remained significant when considering bedrock and mineral or macrobiota content (Fig. 3.2b, Table S3.1). The pH varied from 5.03 to 8.89, with the average pH being

significantly lower for granite samples (5.89 ± 0.49) than for gneiss samples (6.94 ± 1.11) (Welch Two Sample t-test; $P < 0.001$), yet the latter showed a very broad pH range, with sample PB1107 having the overall lowest pH. The pH was also significantly lower when macrobiota were present (Welch Two Sample t-test; $P < 0.005$), and remained so when macrobiota containing samples were further split according to bedrock (Fig. 3.2c, Table S3.2). The electric conductivity ranged from 6 µS cm^{-1} to 890 µS cm^{-1} (on average 99.98 ± 176.83 µS cm^{-1}) but was not significantly different between bedrock type or sample class (macrobiota present versus mineral; Fig 3.2d, Table S3.1). TOC-content showed a weak but significant negative correlation with pH, and a significant moderate positive correlation with both electric conductivity and moisture content (Table 3.2). Electric conductivity and pH were positively correlated ($r = 0.41$, $P < 0.01$).

Table 3.2. Pearson correlation (*r*) and significance levels (*P*) among abiotic variables and between abiotic variables and diversity indices and estimators. Only significant (bold) or nearly significant ($P < 0.1$) effects are displayed.

	TOC		Electric conductivity		pH		Moisture	
	r	*P*	*r*	*P*	*r*	*P*	*r*	*P*
Electric conductivity	**0.46**	**< 0.001**	**-**	**-**	**0.41**	**< 0.005**		
pH	**-0.28**	**< 0.05**	**0.41**	**< 0.005**	**-**	**-**		
Moisture	**0.47**	**< 0.001**					**-**	**-**
OTUs							**0.28**	**< 0.05**
Shannon							**0.36**	**< 0.01**
Inverse Simpson	0.26	0.07					**0.3**	**< 0.05**
Chao1							**0.29**	**< 0.05**

TAXONOMIC COMPOSITION OF THE BACTERIAL COMMUNITIES

The Illumina sequencing resulted in nearly four million reads, of which 395 512 high quality sequences remained for the 52 samples after processing. A total of 3360 OTUs were obtained based on a 97 % cut-off level, which could be assigned to 24 phyla and 189 genera. At the phylum level and genus level, 214 OTUs (985 sequences) and 2534 OTUs (314 419

sequences), respectively, remained unclassified. The average number of sequences per sample equalled 7606 ± 4574 with on average 260 ± 87 OTUs.

In total, 32 % of the reads were identified as belonging to the *Acidobacteria* (Fig. S3.1), of which 95 % were closely related to isolate *Ellin6075* (Joseph 2003) and assigned to class *Chloracidobacteria* (subgroup 4). *Proteobacteria*, *Cyanobacteria* and *Actinobacteria* were the other most abundant phyla (21 – 15 %). *Chloroflexi*, the candidate division *FBP*, *Bacteroidetes* and *Armatimonadetes* represented between 3.3 and 2.6 % of the sequences, while the remaining phyla all represented less than 1 % of the sequences (or less than 3 % of the total number of sequences combined). *Actinobacteria* had the highest number of OTUs (689), followed by the *Proteobacteria* (683), *Bacteroidetes* (397) and *Chloroflexi* (355). However, when the number of OTUs in relation to the number of sequences was considered, *Firmicutes* had the highest relative number of OTUs (44 for 126 sequences) for those phyla represented by more than 12 sequences (i.e., excluding (candidate) phyla *BRC1*, *Elusimicrobia*, *GN02*, *AD3*, *Nitrospira*), followed by candidate phylum OD1 (26 OTUs, 209 sequences) and *Planctomycetes* (176 OTUs, 2149 sequences). Interestingly, *Acidobacteria* were represented by only 145 OTUs.

DRIVERS OF MICROBIAL COMMUNITY STRUCTURE

At the phylum level, *Proteobacteria* and *Acidobacteria* respectively showed a significant positive and negative correlation with TOC, electric conductivity and moisture content (Table 3.3). *Armatimonadetes*, candidate phylum *FBP*, and *Planctomycetes* showed a positive and *Actinobacteria* a negative correlation with TOC and moisture content, while *Armatimonadetes* additionally showed a negative correlation with pH. *Deinococcus-Thermus* showed a positive correlation with electric conductivity, and inversely, a negative correlation

Table 3.3. Pearson correlation (*r*) and significance levels (*P*) between abiotic variables and the relative abundances of the 13 most occurring phyla. Only significant correlations are displayed.

	TOC		Electric conductivity		pH		Moisture	
	r	*P*	*r*	*P*	*r*	*P*	*r*	*P*
Acidobacteria	-0.35	< 0.05	-0.38	< 0.05			-0.32	< 0.05
Actinobacteria	-0.41	< 0.005					-0.49	< 0.001
Armatimonadetes	0.52	< 0.001			-0.3	< 0.05	0.41	< 0.005
Cyanobacteria							0.32	< 0.05
FBP	0.43	< 0.005					0.31	< 0.05
Planctomycetes	0.3	< 0.05					0.29	< 0.05
Proteobacteria	0.58	< 0.001	0.37	< 0.01			0.56	< 0.001
Thermi			0.31	< 0.05			-0.31	< 0.05

with moisture content. Lastly, *Cyanobacteria* showed only a significant positive correlation with moisture content, while no trends were observed for the remaining phyla.

A SIMPROF analysis of the Illumina dataset at the OTU level (Fig. 3.3) and the ARISA dataset (Fig. S3.2) using the Bray-Curtis dissimilarity index showed comparable patterns and revealed a distinct clustering generally coinciding with bedrock type and the presence of macrobiota. A detailed description of the community members in the different significant clusters can be found in the supplementary information. This was confirmed by a Mantel test which showed a significant correlation between the Illumina and ARISA datasets ($P < 0.01$, $r = 0.33$; 9999 permutations; Hellinger transformed data). The granite cluster could be further subdivided into two main sub-clusters based on a Bray-Curtis dissimilarity cut-off of about 0.85, which could be further divided into smaller subclusters based on a dissimilarity cut-off of 0.75 (see Supplementary Information for a more comprehensive description).

Cluster I (Fig. 3.3) showed a relatively high abundance of *Actinobacteria*, with the samples sharing several distinct OTUs (notably most *Intrasporangiaceae*, including genus *Terracoccus* and related unclassified taxa), and a relatively low abundance of *Proteobacteria*

and the near absence of *Cyanobacteria* compared to the other samples within the granite cluster. Granite cluster II contained several distinct acidobacterial *Chloracidobacteria* and *Solibacteres* (including Candidatus *Solibacter*) OTUs, and also the only two *Koribacteraceae* OTUs, one of which was identified as Candidatus *Koribacter*. Furthermore, a relatively large number of *Armatimonadetes* and *Gemmata* (*Planctomycetes*) OTUs were present. Nearly all moss and lichen containing samples were restricted to this cluster. Markedly, five gneiss derived samples belonged to the granite cluster, of which two Perlebandet samples, with sample PB1109 containing lichens. The third Perlebandet sample, PB1107, was the overall outlier, yet still grouped more closely within the granite cluster and, interestingly, also contained moss and lichens. Two other gneiss samples originated from the eastern Duboisbreen region, and a last one from Teltet, which showed the highest overall similarity with an Utsteinen sample.

While samples on gneiss spanned a wider gradient in pH, they appeared to be more similar in terms of OTU composition (Fig. S3.3). They displayed generally higher proportions of *Acidobacteria* and *Actinobacteria*, and lower numbers of *Proteobacteria* and *Cyanobacteria*. One granite sample, KB127, showed notable similarity with the gneiss samples and in the field gneiss was visibly present, suggesting that this sample may at least partly consist of gneiss. The gneiss cluster III displayed a more gradual splitting of its members into small (2-3 sample) groups, but two larger subclusters were present, one of which contained nearly all Brattnipane samples, while the other seemed to be at least partly driven by several *Chloroflexi* OTUs.

A redundancy analysis of both the Illumina and ARISA datasets showed that TOC, followed by pH, electric conductivity (only for the Illumina data), bedrock type and moisture content significantly ($P < 0.001$) explained variation in the community structure (Figures 3.4a and 3.4b), while TC was not significant. An RDA with a subset of 45 samples for which

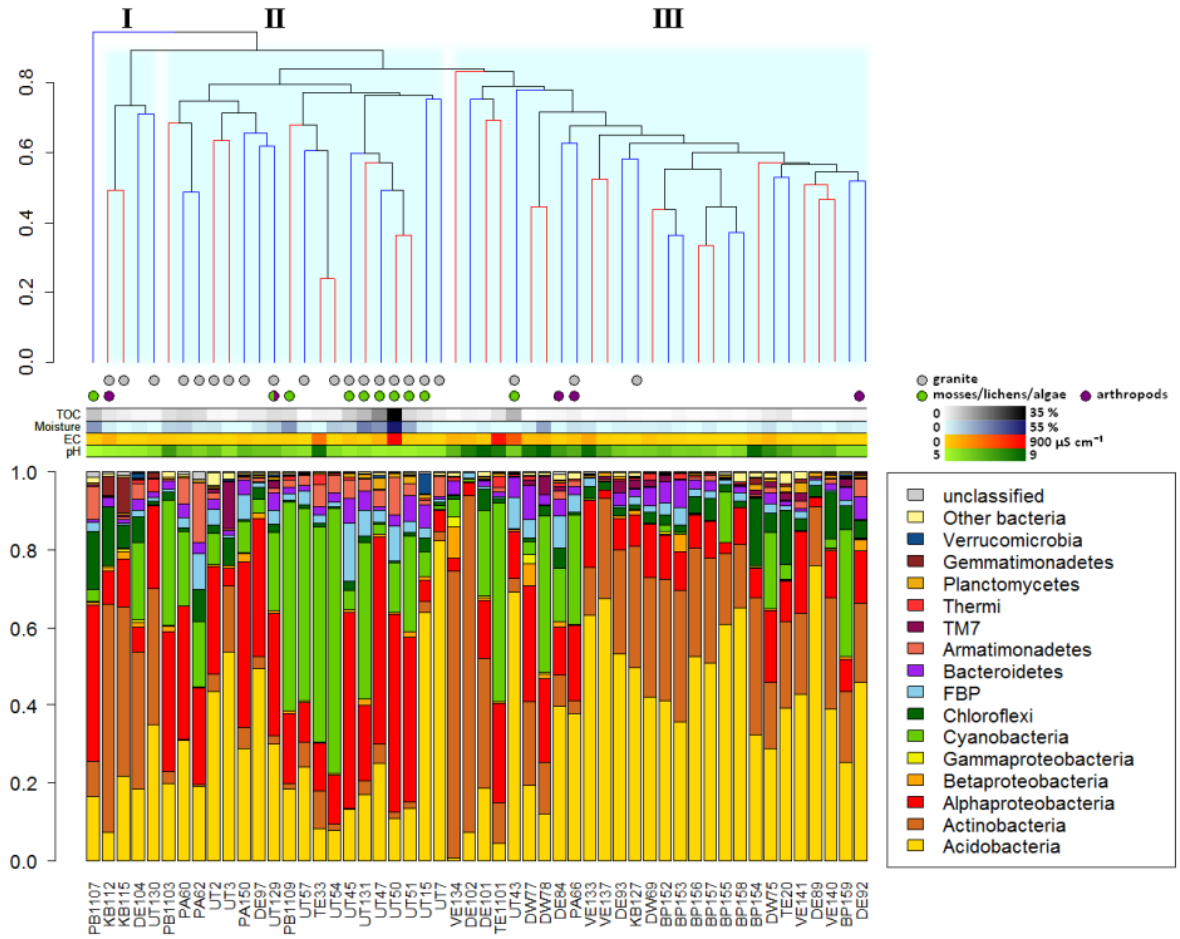

Figure 3.3. The dendrogram at the top shows the significant SIMPROF clusters of samples shown in black, while non-significant clusters are alternating blue or red. The gray circles underneath the dendrogram show the presence of granite (none for gneiss samples), while the green circles show the presence of macrobiota (moss, lichen, algae, or a combination of these), and purple circles arthropods (mites or collembolans). Values of the significant abiotic variables are shown in the bars, with percentages of the total organic carbon (TOC) and moisture content, and the electric conductivity (µS cm^{-1}) and pH. Values of TOC were square rooted to enhance visibility. At the bottom, the barplot shows the relative abundances of the 13 best represented phyla (the *Proteobacteria* are represented by their classes).

elevational data was also present did not show a significant effect of elevation (data not shown). Spatial variables explained a larger part of the variation in the ARISA dataset (five significant PCNM variables) compared to the Illumina dataset (3 variables). The importance of the different abiotic variables in explaining patterns in OTU composition was further confirmed by the variation partitioning analysis. This revealed that ~18 % of the observed variance could be explained by the selected abiotic and spatial variables and that geographical

distance by itself accounted for 3 % of the observed variance, while the abiotic variables and the overlap of the abiotic and spatial variables accounted for 9.4 % and 5.6 % respectively. Bedrock by itself significantly explained 1.5 % of the variance and an additional 4.5 % through the overlap with the other abiotic variables.

Moisture content showed a weak but significant positive correlation with the richness and diversity indices and number of OTUs (Table 3.2). Moreover, the presence of moss or 'black soil' resulted in a significantly higher inverse Simpson value (ANOVA; $P < 0.05$ for both) (data not shown).

3.5. Discussion

FACTORS STRUCTURING BACTERIAL COMMUNITY STRUCTURE

The multivariate analyses of both the Illumina and ARISA datasets were congruent and revealed that total organic carbon and related to this the presence of macrobiota (i.e., algae, mosses and lichens), pH, bedrock type, moisture content, and, additionally for the Illumina dataset, electric conductivity, significantly explained variation in the bacterial communities. The variation partitioning analysis further revealed that 9.4 % of the variation could be uniquely explained by the abiotic variables, 3 % by geographical distance, and 5.6 % by the overlap between both groups of predictors. It follows that environmental factors largely shape the bacterial communities in this part of the Sør Rondane Mountains. These significant variables were also shown to exert a profound influence on bacterial assemblages in terrestrial habitats in other Antarctic ice-free regions. For example, the effects of pH and electric conductivity have been well documented in soils from the Transantarctic Mountains and the McMurdo Dry Valleys (e.g., Magalhães *et al.* 2012; Van Horn *et al.* 2013; Geyer *et al.* 2014; Kim *et al.* 2015), and the influence of moisture content and soil organic matter was also

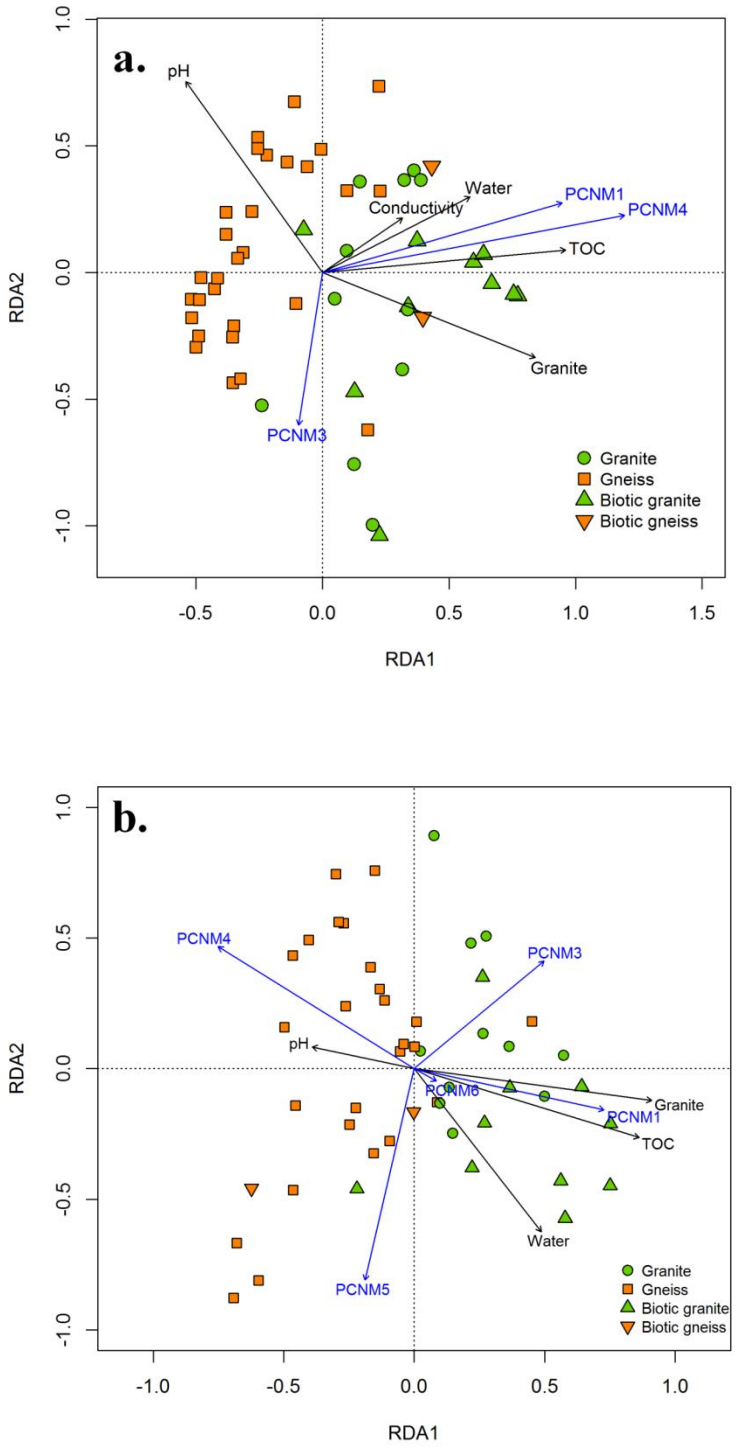

Figure 3.4. Redundancy analysis (RDA) plot of the 52 Illumina sequenced samples (a) and 49 ARISA samples (b). Granite derived samples are represented by green circles (mineral samples) or triangles (macrobiota samples); gneiss derived samples are represented by orange squares (mineral samples) or inversed triangles (macrobiota samples). The spatial variables (PCNMs) are indicated in blue, the environmental factors in black.

demonstrated in some studies from the latter region (Niederberger et al., 2008, 2015; Geyer et al., 2014), while in soils from other areas in the MDV (Wood, Rueckert, et al., 2008), the

Maritime Antarctic (Newsham et al., 2010) and the Transantarctic Mountains (Magalhães et al., 2012), moisture content had either no or only a minor effect on the bacterial community composition or diversity. Bedrock type also significantly explained a portion (1.5 %) of the variation in community structure. Although bedrock type appeared to affect the presence of vegetation, as lichens, mosses and algae were more abundant on granite than on gneiss (Fig. 3.2), the variation explained by bedrock type was independent from that explained by the other environmental factors, including TOC content and hence the presence of macroscopic vegetation cover. However, no significant differences were observed in OTU richness between gneiss and granite samples, which is likely due the mineral composition of both bedrock types being essentially the same. Indeed, preliminary lead (Pb) isotopic analysis of individual feldspars from gravels at several sites, and cosmogenic surface exposure dating analysis of granite-syenite boulders suggested that the gneiss in the Sør Rondane Mountains is likely to be of local origin (S. Roberts, unpublished data), and thus largely has the same chemical composition of the granite. However, the banded texture of gneiss makes it more vulnerable to physical or chemical weathering. Hence, the observed differences in the bacterial communities between the bedrock types could be related to differences in weathering processes, which in turn can lead to small differences in the chemical composition between both granite and gneiss. More in particular, the darker layers with mafic minerals in the gneiss can act as a transport zone of meteoric water, which potentially results in the release of base cations. This can also explain the generally higher pH of gneiss derived soils, which could result in differences in nutrient availabilities that potentially select for different bacterial assemblages between both bedrock types. Interestingly, while communities on gneiss derived soils generally differed from those on granite derived soils, the samples from Perlebandet (gneiss) grouped in a significant cluster composed of granite samples (Fig. 2). This is possibly related to the particular gneiss subtype of Perlebandet (augen gneiss), which

is much harder than the gneiss of the main mountain range and hence more resistant to weathering as similarly observed in the granite derived soils. Moreover, the moisture retention is much higher in the augen gneiss and granite compared with the gneiss of the other regions in the study area, which, combined with lower weathering rates, provides better and more stable conditions for macrobiotic growth. Combined, it can be hypothesized that bedrock type, and related to this weathering rates and the availability of mineral energy sources and moisture, will be an important driver in the oligotrophic Antarctic mineral soils, as was, for example, also demonstrated for oligotrophic cave environments (Barton et al., 2007). Moreover, it has also been shown that rock or soil types with a higher mineral diversity (e.g. siliceous substratum) harbour more diverse bacterial communities (Barton et al., 2007; Meola et al., 2014) and that differences in for example mineral and (trace) element composition, or soil texture, are important factors in shaping bacterial communities in Antarctica and elsewhere on Earth (Barton et al., 2007; Mitchell et al., 2013; Yarwood et al., 2014; Kim et al., 2015).

In addition to bedrock type, the presence of moss or lichen also resulted in very distinct bacterial communities (Fig. S3.3). This will, on the one hand, be due to an increased availability of organic nutrients (debris, exudates) resulting in a larger proportion of heterotrophs, but also due to differences in microclimatic conditions, such as a reduced wind exposure and an increased humidity and heat retention (e.g. Cannone and Guglielmin 2009). Moreover, moss and lichen containing samples in soils from both bedrock types shared a relatively high number of OTUs (Fig. S3.4), compared to those shared with their respective abiotic soils, underscoring the importance of vegetation cover for bacterial community structure. However, the number of taxa above OTU level was comparable between macrobiotic samples and mineral soils, although, for example, relatively more *Proteobacteria* lineages were present in the former. Interestingly, while *Firmicutes* were found to be the most

dominant phylum in the rhizosphere of vascular Antarctic plants (Teixeira et al., 2010) and highly abundant in MDV samples with a higher moisture content (Van Horn et al., 2013), they were nearly absent from our dataset (only 126 sequences), and mainly restricted to two mineral gneiss samples. Similarly, relatively few *Deinococcus-Thermus* were recovered from the soils without vegetation (2233 sequences, of which 1156 identified as *Truepera*), while it could be expected that because of their resistance to abiotic stresses (Battista, 1997), they should be able to resist the harsh terrestrial conditions as for example observed in dry mineral soils from Victoria Land (Niederberger et al., 2008; Pointing et al., 2009b). Also vegetation cover and bedrock type appeared to have a combined effect on two out of the eight most abundant phyla (*Bacterioidetes* and *Planctomycetes*; Table 3.1). Although we are aware that the number of samples with vegetation on gneiss is lower than that on granite, leading to an imbalanced sampling design, the effect of vegetation on soil properties, which in turn affects bacterial community structure, is in agreement with what can be expected. For example, the combined effect of both factors might be due to the influence of vegetation on e.g. the pH or the tendency of the soil to store moisture. By contrast, while vegetation cover appears to be important, OTU richness was unrelated to TOC content (Table 3.2). Interestingly, similar conclusions were drawn from a clone library survey from ornithogenic soils in the MDV (Aislabie et al., 2009), while Geyer *et al.* (2013) found a productivity-diversity relation in a variety of MDV soils using T-RFLP. We did, however, notice a significant trend of increasing OTU-richness with increasing moisture content, and Shannon and inverse Simpson diversity indices were significantly higher in moister samples too. This is in disagreement to what Newsham, Pearce and Bridge (2010) as well as Stomeo *et al.* (2012) found in Mars Oasis (West Antarctic Peninsula) and MDV soils, respectively. The importance of the moisture content in the soils in the Sør Rondane Mountains in contrast to Mars Oasis and the MDV might be related to the wide gradient covered in the present study.

In addition to the abovementioned environmental variables, spatial distance appeared to play a significant, yet relatively minor role in our study sites; three out of the six PCNM variables significantly explained only 3 % of the observed variation in the Illumina dataset. Variation partitioning further revealed that the overlap between geographic distance with abiotic variables was more important than distance itself, which was also observed in the MDV and the Transantarctic Mountains (Geyer et al., 2013; Sokol et al., 2013; Chong et al., 2015). Nevertheless, some spatial autocorrelation is apparent since samples in close proximity tend to be more alike, which was clearly evident in the samples from Brattnipane. More striking, however, is the fact that the communities in the (northern) nunataks differed substantially from those situated on similar bedrock samples in the main mountain range to the South. This suggests limited dispersal capacities, most likely in combination with microclimatic differences as also noted in other areas from the Antarctic (see Chong, Pearce and Convey 2015 for a review). Indeed, the different regions of the main mountain range showed a relatively high similarity, which may be partly linked to their closer geographic connectedness, possibly facilitating dispersal of organisms over short distances, in combination with the severe influence of the nearby polar plateau.

BACTERIAL COMMUNITY COMPOSITION

In general, the main phyla observed are globally dominant in soils (Janssen, 2006) and also more or less the same as those reported from other Antarctic terrestrial studies (Yergeau, Newsham, et al., 2007; Pointing et al., 2009b; Teixeira et al., 2010; Van Horn et al., 2013; Bakermans et al., 2014; Kim et al., 2015), namely *Proteobacteria*, *Actinobacteria*, *Acidobacteria*, *Chloroflexi*, *Verrucomicrobia*, *Bacteroidetes*, *Armatimonadetes*, *Gemmatimonadetes* and *Planctomycetes*, and additionally candidate division *FBP* and *Cyanobacteria* (Table 3.1, Fig. S3.1). However, half of our samples were dominated by *Acidobacteria*. This was particularly the case in gneiss derived soils with a very low TOC

concentration, which might be related to the slow growth rate of *Acidobacteria* which usually thrive in oligotrophic conditions (Fierer et al., 2007). Nearly all acidobacterial reads were assigned to the class *Chloracidobacteria* (Subdivision 4 of the *Acidobacteria*), which was represented by 119 288 sequences and constituted ~95 % of all *Acidobacteria* sequences. The phylum *Acidobacteria* as a whole displayed a relatively low diversity (only 145 OTUs) while having 32 % of the total reads. Similar high relative abundances of *Acidobacteria* have recently been reported from Victoria Land, where they appeared to be associated with the relatively richer – albeit still oligotrophic – soils (Van Horn *et al.* 2013), characterised by a higher pH (Van Horn et al., 2013; Kim et al., 2015). Among the *Acidobacteria, Chloracidobacteria* are considered to be aerobic anoxygenic photoheterotrophs, which contain distinctive bacteriochlorophylls (Bchl *a* and *c*) (Bryant et al., 2007). Although we cannot confirm the actual presence of bacteriochlorophylls *a* and *c* in our samples, the ability to use solar energy in the Antarctic as a means to generate ATP would be highly advantageous (Dana et al., 1998), and might indeed lead to their high relative abundance in many of our samples. Tahon *et al*. (2016) found evidence for a large diversity of light capturing *pufM* genes in some of these samples, at least suggesting the possibility of exploiting this energy source by soil organisms. In the SRM, contrary to the *Chloracidobacteria*, OTUs from the *Acidobacteriaceae* (genus *Terriglobus* and unclassifieds) were restricted to moss-containing samples, while *Koribacteraceae* were nearly completely restricted to granite samples with macrobiota. Members of the *Solibacterales* showed an increased relative abundance on granite derived soils with relatively high carbon content. It is known that Candidatus *Solibacter* is a chemoorganotroph that is able to use complex organic carbon sources such as cellulose, and stabilizes its environment through biofilm production (Ward et al., 2009). Candidatus *Solibacter* was also the most recovered genus from a

metagenomics study of Mars Oasis, a biodiversity hotspot in Maritime Antarctica (Pearce et al., 2012).

In addition to *Acidobacteria*, *Cyanobacteria* were also widespread (75 % of the samples). However they showed a relatively low abundance compared with aquatic habitats (Tytgat et al., 2014). Their presence has however important ecological consequences, because while nitrogen is thought to be a major limiting factor in Antarctic terrestrial systems, it is generally assumed that N-fixation in Antarctica is mainly performed by *Cyanobacteria* (Cowan, Sohm, et al., 2011), either free-living or as cyanobionts in lichens (Seneviratne and Indrasena, 2006; Yergeau, Kang, et al., 2007). It is also assumed that other free-living organisms rarely fix nitrogen in these environments due to energetic constraints (Raymond et al., 2004). However, also within the *Actinobacteria*, certain species (e.g., *Arthrobacter*, *Frankia*, *Rothia* and *Corynebacterium*) have N-fixation potential (Sellstedt and Richau, 2013). Considering the relatively high presence of *Actinobacteria* in our soils, and that *Arthrobacter* and unclassified *Frankiaceae* related OTUs were among the most represented *Actinobacteria*, these taxa can potentially be important N-fixators in addition to *Cyanobacteria* in these terrestrial environments. In the *Cyanobacteria* it is hypothesized that in particular taxa belonging to the *Nostocales* order play a major role in nitrogen fixing in similar ecosystems (Cowan, Sohm, et al., 2011). Indeed, a clone library study by Tahon *et al*. (2016) from the Utsteinen area showed that only *nifH* sequences from *Nostoc*-related organisms were recovered. In our dataset, an OTU associated with the *Nostocales* showing 99.1 % similarity to the recently described heterocystous genus *Toxopsis* (Lamprinou et al., 2012) was the fifth most relative abundant (11 211 sequences) OTU (after four *Chloracidobacteria* OTUs). However, it was recovered from only 19 samples and seemed closely associated with the presence of moss, lichens, a high TOC content and granite soils. Overall, ten cyanobacterial OTUs were among the 30 most abundant OTUs and appeared to be closely related to either *Nostoc*, several

Phormidium spp., *Microcoleus vaginatus* CJ1-U2-KK1, and *Crinalium* PCC9333. In general, the high number of unclassified sequences, which underscores the need for additional biodiversity assessments in Antarctic inland locations as also revealed by other studies of different habitats in the regions (Obbels et al. in press, this issue) and other regions (Chong et al., 2013).

3.6. Conclusions

In conclusion, *Acidobacteria* (*Chloracidobacteria*) and *Actinobacteria* (*Actinomycetales*) dominated mineral soil samples situated on gneiss derived bedrock, while *Proteobacteria* (*Sphingomonadaceae*), *Cyanobacteria*, *Armatimonadetes* and candidate division *FBP* (OP11) dominated soil samples with a high total organic carbon content mainly on granite derived bedrock. It follows that the structure of soil microbial communities in the Sør Rondane Mountains is mainly controlled by (i) the availability of moisture and TOC, which is related to the presence of mosses or lichens, and (ii) bedrock type, which most likely influences pH, electric conductivity and substrate stability through differences in weathering rates.

3.7. Acknowledgements

We are grateful to Veerle Vandenhende (Laboratory of Soil Science, UGent) for performing the carbon measurements and Wim Van Roy (Geography Department, UGent) for providing the SRM maps.

3.8. Supplementary Information

Table S3.1. Results of a Tukey post-hoc test of an ANOVA, showing the significant or nearly significant differences between four classes of samples (mineral gneiss, mineral granite, gneiss with macrobiota, granite with macrobiota).

	Gnb-Gn	Gr-Gn	Grb-Gn	Gr-Gnb	Grb-Gnb	Grb-Gr
	P	*P*	*P*	*P*	*P*	*P*
pH	< 0.1	< 0.01	< 0.01			
Electric conductivity						< 0.1
TOC	< 0.01	< 0.1	< 0.001			< 0.001
Chao1						
Moisture content	< 0.05		< 0.01	< 0.05		< 0.05
Actinobacteria			< 0.01			
Armatimonadetes						
Bacteroidetes			< 0.05			
FBP			< 0.001			< 0.05
Gemmatimonadetes						
Planctomycetes			< 0.05			< 0.001
Proteobacteria			< 0.01			< 0.05
Verrucomicrobia			< 0.1			< 0.1

Gn: mineral gneiss, n = 29; Gnb: macrobiota containing gneiss, n = 2; Gr: mineral granite, n = 13; Grb: macrobiota containing granite, n = 8.
TOC: total organic carbon

A more detailed description of the clusters and subclusters (Fig. 3.3) is given below.

Cluster I showed two smaller clusters with cluster Ia consisting of two moderately similar (~ Bray-Curtis value of ~0.5) Ketelersbreen samples containing a relatively high abundance of *Gemmatimonadetes* order *N1423WL* (~5-10% of reads) and *Chloroflexi* (*Ktedonobacteria*) and several unclassified *Actinomycetales* (*Actinobacteria*) and *Acetobacteraceae* (*Alphaproteobacteria*) OTUs, not shared with cluster Ib. In one of these samples, KB112, collembolans were found. Despite the in general higher similarity between samples of the same bedrock type, those from Ketelersbreen formed an exception in the granite cluster, showing a relatively low similarity with both other granite derived samples (the Utsteinen and

Pingvinane nunataks, situated 15 km to the north and north west, respectively), as well as the adjoining gneiss derived samples. These southernmost samples were situated on a granite outcrop which is surrounded by gneiss in the main mountain range.

Cluster Ib was characterized by several shared *Acetobacteraceae* and *Actinomycetales* (*Sporichthyaceae* and *Nocardioidaceae*) OTUs.

The granite cluster II samples shared several distinct acidobacterial *Chloracidobacteria* and *Solibacteres* (including Candidatus *Solibacter*) OTUs, and also the only two *Koribacteraceae* OTUs, one of which was identified as Candidatus *Koribacter*. Furthermore, a relatively large number of *Armatimonadetes* and candidate phylum *FBP* were present, and *Gemmata* (*Planctomycetes*) OTUs were shared. Nearly all moss and lichen containing samples were restricted to this cluster.

Cluster II could be subdivided into five smaller clusters. Subcluster IIa consisted of three samples, two from Pingvinane and one from more northwestern Perlebandet. Analysis showed they had a higher TOC content. Interestingly, black soil was reported to be present in all three samples. They had more or less similar ratios of *Acidobacteria*, *Proteobacteria* and *Cyanobacteria*, while the PA samples had additionally higher ratios of *Armatimonadetes* and *FBP*. Nevertheless, clustering seemed to be mainly driven by the presence of several *Alphaproteobacteria* OTUs, notably from the genus *Zymomonas*. Cluster IIb showed a generally elevated *Acidobacteria* presence, but otherwise lacked any clear distinguishing features, and samples were nearly equally highly dissimilar. Sample UT129 however is characterized by the presence of moss, lichens and both collembolans and mites. Cluster IIc was not only dominated by *Cyanobacteria*, but samples in this cluster also showed a broad taxonomic diversity, represented by in particular the genera *Pseudoanabaena*, *Leptolyngbya*, *Phormidium* and unclassified *Nostacaceae* and *Xenococcaceae*, and additionally shared

nearly all sequences of a single *Chamaesiphonaceae* OTU. Interestingly, two samples were gneiss derived, and two were granite derived. Samples TE33 and UT54 showed the highest similarity of all samples, despite having different bedrock types and being located nearly 7 km apart, while sample UT57 was only 45 m away from sample UT54 and sharing the same type of bedrock. Cluster IId consisted of only samples associated with macrobiota (moss and/or lichen). All had relatively high levels of TOC and water, and a low pH. *Alphaproteobacteria* generally dominated this cluster, with various numbers of *Cyanobacteria*. Together, *FBP*, *Armatimonadetes* and *Bacteroidetes* represented a relatively high proportion compared to other clusters. *Actinobacteria* had very low abundances and were absent from sample UT45. OTUs representing a.o. Candidatus *Solibacter* (*Acidobacteria*) and several *Armatimonadetes* orders were characteristic for this cluster. The samples within the last granite subcluster (IIe) were highly dissimilar but shared a very high relative abundance of *Acidobacteria*, which could virtually all be attributed to four chloracidobacterial *Ellin6075*-related OTUs, with some additional *Solibacterales* OTUs (especially the candidate genus *Solibacter*). UT15 furthermore showed equal numbers of *Alphaproteobacteria*, *Cyanobacteria*, *Bacteroidetes* and, noteworthy, *Chthoniobacteriaceae* (*Verrucomicrobia*) OTUs related to grassland clone DA101 (Felske and Akkermans, 1998), with this latter phylum largely being restricted to the granite cluster. In contrast, sample UT7 contained about 5% reads of *Armatimonadetes* and *Alphaproteobacteria* each, the former mainly represented by order *FW68*, and the latter by *Acetobacteraceae* and *Sphingomonadaceae* related OTUs.

Interestingly, gneiss derived samples PB1109, which contained lichens, and PB1103, TE33 and DE104 clustered with the granite samples, while PB1107, with both lichens and moss, formed the overall outgroup, but still clustered more closely with the granite derived samples. The Perlebandet nunataks appeared to be slightly more similar to the relatively nearby

Utsteinen and Pingvinane nunataks (15-20 km distance), despite having a different substrate type. For example, all Perlebandet samples lacked an *Actinobacteria* family *AKIW874* (order *Acidimicrobiales*) OTU which was found in nearly every other gneiss derived sample. Nevertheless, PB1107 differed substantially from the other PB samples, too. Despite moss and lichens being present in this sample, the prokaryotic community showed only a moderate similarity with the samples of macrobiota situated on granite, sharing e.g. the *Terriglobus* and other *Acidobacteriaceae* OTUs, a Candidatus *Solibacter* OTU and other members of the *Solibacterales*, and several alphaproteobacterial groups. It showed a relatively high proportion of OTUs related to the *Chloroflexi* class *Ktedonobacteria* and a number of unclassified *Actinomycetales* OTUs largely restricted to this sample, accounting for nearly a third of the actinobacterial reads in this sample.

The other gneiss derived samples formed a distinct cluster and generally had a higher proportion of *Acidobacteria* and *Actinobacteria*, and lower numbers of *Proteobacteria* and *Cyanobacteria*. Samples within the gneiss cluster hence showed relatively higher similarities compared to the granite derived samples (see also supplementary figure S3). Sample VE134, which contained 74 % actinobacterial reads, relatively few *Alphaproteobacteria* but more *Betaproteobacteria*, formed an outlier within this cluster. Next, three highly dissimilar samples (DE102, DE101 and TE1101) split off. DE102 was also dominated by *Actinobacteria*, while TE1101 had a high proportion of *Cyanobacteria* and *Alphaproteobacteria*. Granite derived sample UT43, forming a separate branch, had a large number of acidobacterial sequences, and also a relatively large fraction of *FBP* members (one third of all *FBP* OTUs were present). Interestingly, macroscopic algae were observed in this sample, while *Cyanobacteria* were virtually absent (4 sequences, 3 OTUs). It also had a high TOC content and a high conductivity. The other samples further clustered according to a generally increasing proportion of *Acidobacteria* and decreasing number of *Cyanobacteria*, with most

of the Brattnipane samples forming a distinct cluster. Besides UT43, only two granite derived samples clustered within the gneiss derived samples. KB127 was very dissimilar to the other KB samples. It was taken at the edge of the Ketelersbreen granite outcrop, and contains a fraction of gneiss. Granite sample PA66, containing mites and collembolans, clustered with DE84, which also contained mites.

3.9. Supplementary figures

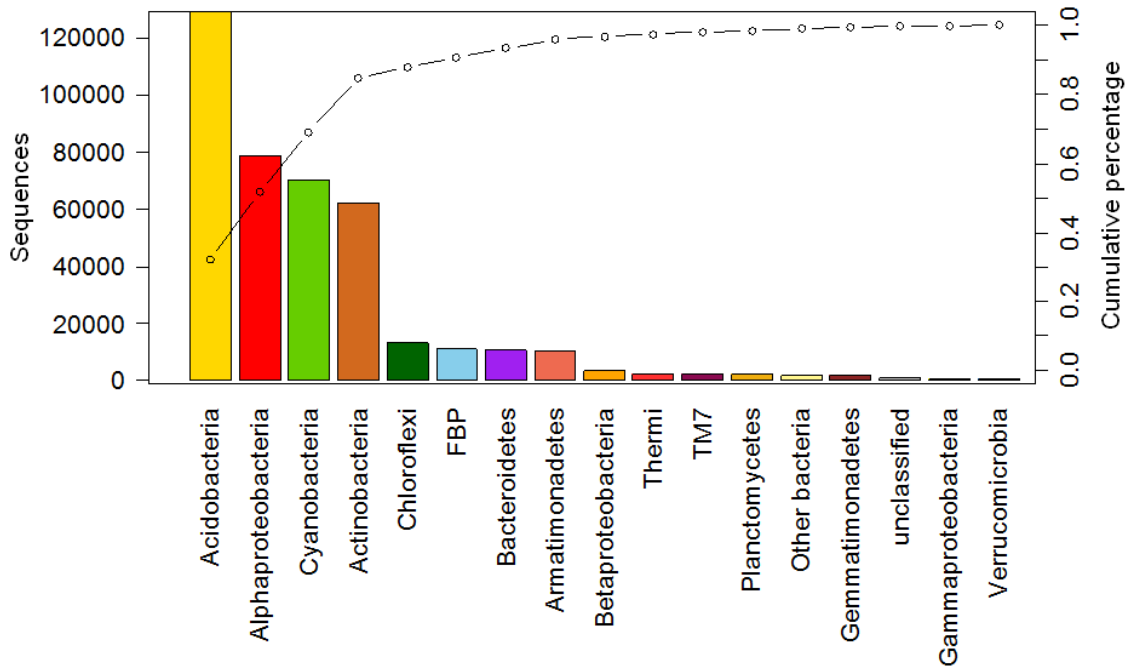

Figure S3.1. Pareto chart of the 13 most abundant phyla, showing the total number of sequences (left y-axis) and their cumulative percentage (right y-axis). The phylum *Proteobacteria* was split into its classes. The bin 'Other bacteria' contains (candidate) phyla *AD3*, *Nitrospirae*, *Elusimicrobia*, *GN02*, *BRC1*, *Fusobacteria*, *Firmicutes*, *OD1*, *Chlorobi*, *WPS-2*, *MVP-21*,and the remaining *Proteobacteria* classes *Epsilonproteobacteria*, *Deltaproteobacteria* and *TA18*.

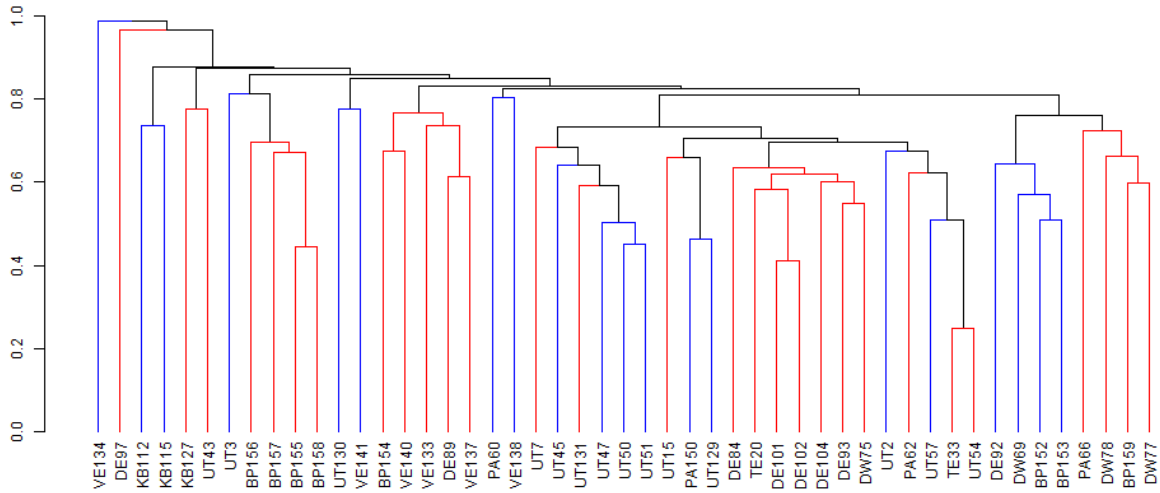

Figure S3.2. SIMPROF analysis of the ARISA dataset. Significant SIMPROF clusters are shown in black, while non-significant clusters are alternating blue or red.

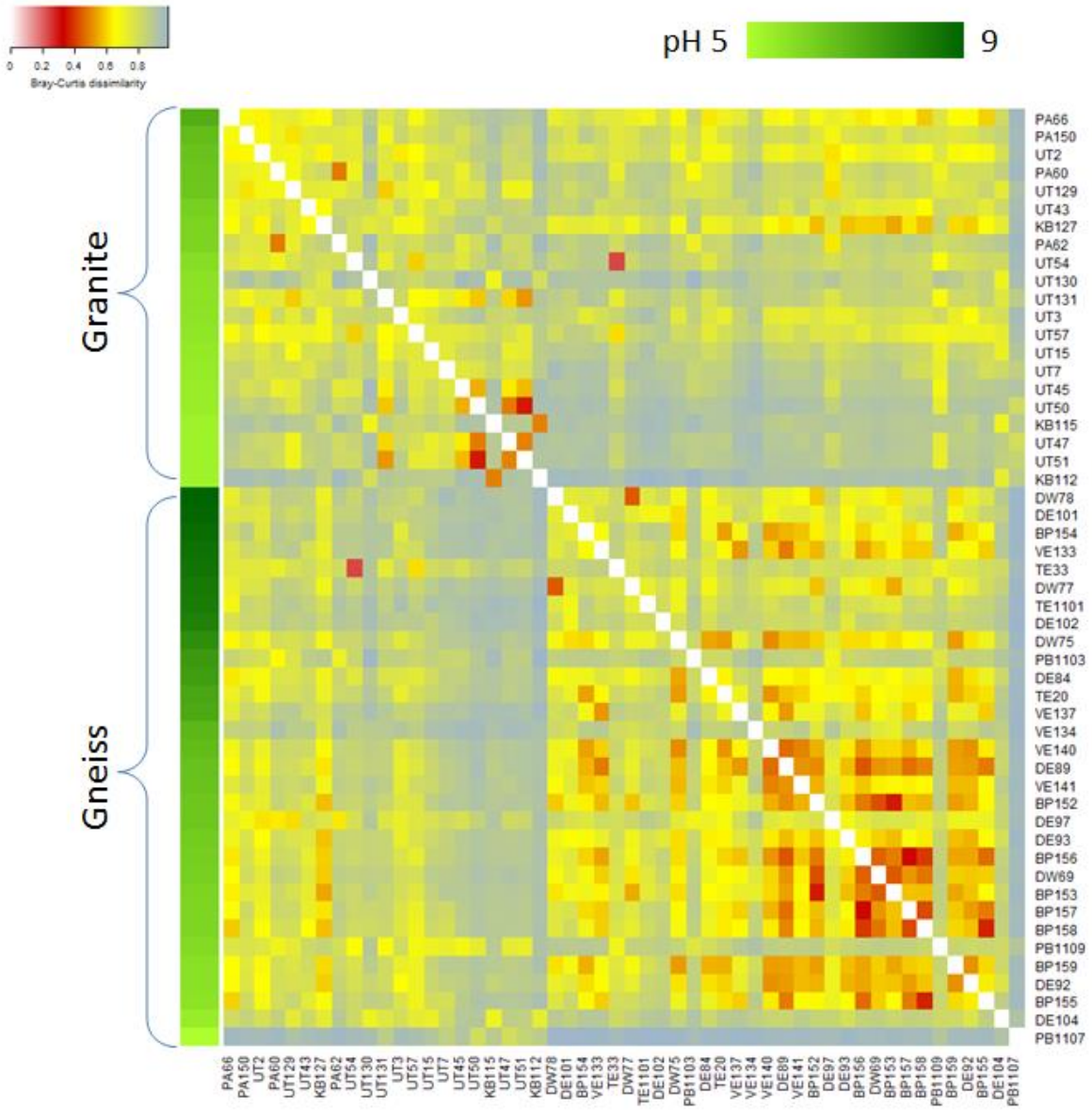

Figure S3.3. Heatmap showing the Bray-Curtis dissimilarities between samples. Samples are ordered according to bedrock and decreasing pH.

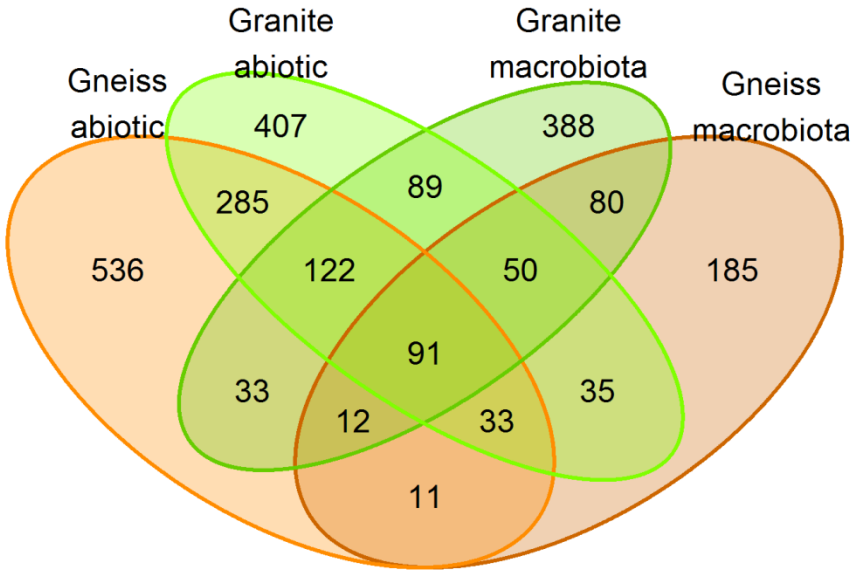

Figure S3.4. Venn diagram showing the number of OTUs shared between the different sample classes (i.e., abiotic or macrobiota containing gneiss or granite samples). OTUs were determined after rarefying to the total number of sequences of the smallest class (i.e., 33 117 sequences in the macrobiotic gneiss samples). The total number of OTUs after rarefying were 1123, 497, 1112 and 865 for classes ‘gneiss abiotic’, ’gneiss macrobiota’, ‘granite abiotic’ and ‘granite macrobiota’ respectively.

Chapter 4. Biogeographical and macroecological patterns in lacustrine benthic microbial mats of the Antarctic Region

Manuscript in preparation
Tytgat, B., Sweetlove, M., Verleyen, E., D'hondt, S., Sabbe, K., Vyverman, W. and Willems, A. Biogeographical and macroecological patterns in lacustrine benthic microbial mats of the Antarctic Region.

Author contributions
BT, EV, WV, KS, MS and AW developed the study. BT and SD processed the samples. BT, MS and EV analysed the data. BT, EV and AW wrote the paper.

4.1. Summary

Benthic microbial mat communities dominate primary production and nutrient cycling in Antarctic lakes. Surveys of their diversity are scarce and have been geographically restricted to a limited number of lakes in the vicinity of research stations. While this revealed the presence of a large diversity in microorganisms and provided information on which environmental parameters structure their community structure, knowledge of biogeographical and macroecological patterns in Antarctic microbial organisms are rare and if available, largely based on the comparative studies, often using different techniques. Here, we analysed a total of 138 samples of benthic microbial communities from eight regions in Antarctica and two Sub-Antarctic islands using a standardized amplicon-based high-throughput approach. OTU-richness significantly increased with decreasing latitude between 85 and 54° S. While a significant difference in community structure between Sub-Antarctica and Antarctica was observed, largely due to mean annual temperature, electric conductivity appeared to be important in explaining turnover patterns between lakes on the Continent additionally to geographic proximity. A high overall diversity was observed. *Cyanobacteria* and *Alphaproteobacteria* appeared to dominate most freshwater samples, while *Bacteroidetes* and *Rhodobacteraceae* (*Alphaproteobacteria*) were more important in high-conductivity lakes. Marion Island was dominated by *Janthinobacterium* (*Betaprotobacteria*) related OTUs.

We conclude that while a biogeographic zoning between the Sub-Antarctic islands and Antarctica is present, possibly related to differences in catchment characteristics and mean annual temperature, differences in bacterial communities on the continent are to a large extent related to specific conductance additional to geographic connectivity.

4.2. Introduction

Recent advances in the knowledge of patterns of the biogeography in terrestrial macro-eukaryotic organisms has revolutionised our insights in the controls and history of life in the Antarctic Region (AR) (Convey et al., 2014; Chong et al., 2015). More in particular, these data revealed a high degree of endemism and strong bioregionalisation patterns in nearly all groups studied, suggesting considerable dispersal limitation in combination with a long evolutionary history on the continent (Convey et al., 2008; Terauds et al., 2012). For example, the traditional view that the AR can be subdivided into three biogeographic regions, namely Continental Antarctica, Maritime Antarctica and the Sub-Antarctic islands was challenged by recent studies that revealed a higher amount of biogeographic provincialism, which can only be explained by the survival of taxa in glacial refugia during Neogene and Pleistocene glacial maxima (Terauds et al., 2012; Convey et al., 2014). However, despite being fundamental to the structure and functioning of Antarctic terrestrial communities (Wynn-Williams, 1996; Tindall, 2004), far less is known about the diversity and distribution of microorganisms in the AR (Chong et al., 2015). What is known comes from a fairly limited amount of inventories of soils and lakes, and from a restricted number of areas (Laybourn-Parry and Wadham, 2014).

Interestingly, the lake ecosystems appear to be true biodiversity hotspots for microorganisms and small invertebrates (Convey et al. 2014). These ecosystems can be found in high latitude ice-free oases, including nunataks in the Transantarctic Mountains (Hodgson et al. 2010). The lakes exhibit a wide diversity of physical and chemical characteristics, ranging from extremely low conductance meltwater ponds to hypersaline lakes, with salinities reaching ten times that of oceanic water (Laybourn-Parry and Pearce, 2007; Dickson et al., 2013; Laybourn-Parry and Wadham, 2014). The lakes can be permanently ice-covered to perennially ice-free, and completely mixed to seasonally stratified or meromictic (Wilkins et

al., 2013). Moreover, lakes in close proximity can differ substantially in their physical and chemical characteristics. It follows that in these ice-free oases a large variety of lacustrine habitats and niches can be potentially present (Wilkins et al., 2013). Despite their generally oligotrophic to ultra-oligotrophic nutrient status, lakes are important sources of carbon in terrestrial Antarctic ecosystems given that they harbour many autotrophs, which can show substantial primary production (Laybourn-Parry, 2002). The vast majority of this primary production is mainly associated with the benthic microbial mats which are generally dominated by filamentous *Cyanobacteria*, most prominently of the order *Oscillatoriales* (Vincent et al., 2000; Cowan and Ah Tow, 2004; de los Ríos et al., 2015). Air bubble formation as a result of photosynthesis, and the movement of melting lake ice due to wind stress and the subsequent mechanical disturbance of the microbial mats, can lead to the detachment of these communities and dispersal via wind into the surrounding terrestrial systems as well as other water bodies. In addition, lake level lowering and the subsequent drying out of these mats can provide important carbon sources to the surrounding soils (Wood, Rueckert, et al., 2008; Cary et al., 2010). The secretion of mucilaginous organic compounds by the filamentous cyanobacteria (exopolysaccharides or extracellular polymeric substances) which bind sediment particles, further leads to stable and buffered habitats, which sustain a diverse consortium of microbial eukaryotic and prokaryotic autotrophs and heterotrophs, as well as small invertebrates (Vincent et al., 2000; Varin et al., 2010; Guerrero and Berlanga, 2013).

Although microbial diversity in aquatic environments is higher than in surrounding terrestrial locations, lacustrine food chains are truncated, too. Freshwater community structure in the maritime Antarctic includes single copepod and anostracan herbivores and a predatory copepod, in addition to benthic cladocerans and ostracods, and microscopic groups. Aquatic faunas are more reduced in Continental Antarctica still (Convey et al., 2008), with

heterotrophic protists generally constituting the top level, while in some coastal lakes, a single crustacean species occurs as top predator (Wilkins et al., 2013). Viral loads, up to an order of magnitude higher than bacteria (Laybourn-Parry and Wadham, 2014), are in the absence of high level predators or grazers thought to be important in the top-down control of the microbial community, and hence a major contributor to the lake carbon cycle.

In this chapter, we aimed to study the diversity and community structure of lacustrine microbial mat associated bacteria along geographical and environmental gradients in regions in continental and Maritime Antarctica, and two Sub-Antarctic Islands, namely Marion Island (Southern Indian Ocean Province) and Macquarie Island (Southern Pacific Ocean Province) ranging in latitude from 85° S to 45° S.

4.3. Material and Methods

Sample collection and selection

Littoral sediment and benthic microbial mat samples were obtained during different sampling campaigns between the austral summers of 1993 and 2013. Material was collected in sterile plastic bags or containers, and stored at -20 °C until processing in the laboratory. In total, 139 samples were selected along a broad geographical and climatic gradient. This included samples from the North-East and the North-West Peninsula, the Transantarctic Mountains (TAM), Dronning Maud Land, Enderby Land, East Antarctica and Victoria Land. We also included samples from Marion Island (MI) in the Southern Indian Ocean Province, and Macquarie Island in the Southern Pacific Ocean Province (Fig. 4.1 and Table 4.1). The mainland samples covered eight of the in total 15 different previously defined Antarctic Conservation Biogeographic Regions (ACBRs; Terauds et al., 2012), and a wide

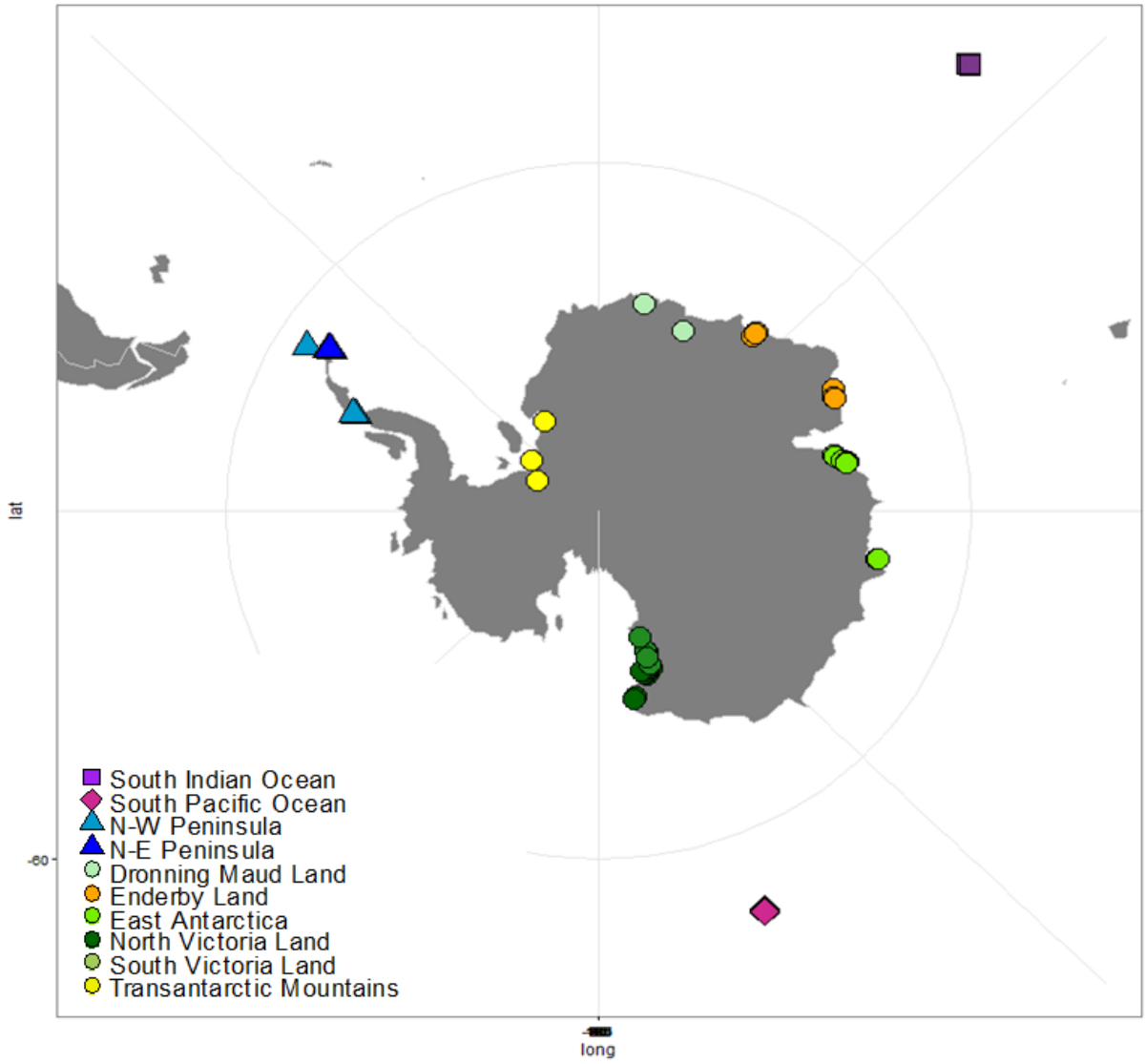

Figure 4.1. Sample locations within the Antarctic and Sub-Antarctic. Colour coding for the regions is constant throughout this chapter. Antarctic regions correspond to the ACBRs recognized by Terauds et al. (2012).

range in conductivity and pH. These factors were previously shown to be important in structuring Antarctic microbial communities (Verleyen et al., 2003, 2010; Chong et al., 2012). Most samples (n = 90) came from freshwater lakes (EC ≤ 1.5 mS/cm), while 9 samples had an electric conductivity of more than 60 mS/cm. No conductivity data was available for 11 samples.

Table 4.1. Overview of the regions included in this study according to Terauds et al. (2012), including total number of samples, and the number of samples in the sampled oases within a region. For some oases (marked by a superscript letter) a reference paper is available with the original descriptions of the samples.

ACBR Region Name	**no.**[1]	**samples**	**Remarks**[2]
North-East Antarctic Peninsula	1	19	Beak Isl. (5), View Point Isl. (3), Vega Isl. (11)
North-West Antarctic Peninsula	3	7	Horseshoe Isl. (2), Pourqoi-Pas Isl. (2), King George Isl. (3)
Enderby Land	5	27	West Ongul Isl.[c] (9), East Ongul Isl.[c] (4), Langhovde[c] (7), Syowa[c] (3), Mawson area (4)
Dronning Maud Land	6	8	Schirmacher Oasis (7), Sør Rondane Mountains (1)
East Antarctica	7	23	Bunger Hills (3), Vestfold Hills (6), Larsemann Hills[a] (9), Bøllingen Isl.[a] (3), Rauer Isls.[a] (2)
North Victoria Land[b]	8	17	
South Victoria Land	9	8	
Transantarctic Mountains	10	8	Forlidas Valley (2), Mount Lowry (1), Davis Valley (3), Shackleton Range (2)
Marion Island	*NA*	19	
Macquarie Island	*NA*	13	

[1]ACBR numbering according to Terauds et al. (2012).
[2]Number of samples in specific oases
[a]Hodgson et al. (2001); Sabbe et al. (2003)
[b]Borghini et al. (2007)
[c]Tavernier et al. (2014)

Ancillary data gathering

Data on temperature was obtained from the CRUTEM4 database (version 7 October 2014; Jones et al., 2012; Osborn and Jones, 2014). Temperature records for the nearest available station were used, and averaged over the period 1993-2013. No temperature data was available for the Transantarctic Mountains, as the average distance to the nearest weather station was considerably more than 500 km. In addition, we also calculated the altitude and approximate shortest distance to the sea using Google earth (15 February 2015; Google Inc.).

DNA extraction and amplification

Per sample, 1 g was used for DNA extraction. First, extracellular DNA was removed following Corinaldesi et al. (2005). DNA extraction was performed according to Zwart (1998). PCRs targeted the V1–V3 hypervariable regions of 16S rRNA genes and were performed in duplicate. Primers pA (AGAGTTTGATCCTGGCTCAG, positions 8–27) (Edwards et al., 1989) and BKL1 (GTATTACCGCGGCTGCTGGCA, positions 536–516) (Cleenwerck et al., 2007) were modified with adapters to complement the Nextera XT index kit (Illumina, USA), and HPLC purified (IDT, Belgium). Each reaction consisted of 2.5 µl 10x buffer (High Fidelity PCR system, Roche, Switzerland), 2.5 µl dNTPs (10 mM) (Life Technologies, Belgium), 0.5 µl of each primer, 0.1 unit of High Fidelity Hot Start polymerase (Roche, Switzerland), 0.5 to 2 µl of DNA template, and was adjusted to a final volume of 25 µl using sterile HPLC water. An initial denaturation step of 3 min at 95 °C was followed by 27 cycles of 30 s at 95 °C, 45 s at 55 °C, and an extended elongation of 3 min at 72 °C in order to try and reduce the number of chimeric sequences (Engelbrektson et al., 2010). A final elongation of 10 min at 72 °C was performed. PCR products were pooled and purified using a slightly modified Ampure beads XT (Agencourt, Beckman Coulter, USA) protocol, where only 0.8 volumes of beads were used and DNA was resuspended in milliQ water.

Library preparation

Nextera XT indices were attached in an eight-cycle version of the amplicon PCR with the indices replacing the primers. PCR products were purified as described above, with DNA resuspended in a 0.10 M 8.5 pH Tris buffer. DNA integrity was checked using a Bio Analyzer (Agilent Technologies, USA) and concentrations were measured using the High Sensitivity array for Qubit (Thermo Fisher Scientific, USA), after which the samples were pooled equimolarly. Amplicons were sequenced on an Illumina MiSeq platform giving 2x300 bp

paired end reads.

Bioinformatics data processing

Paired end reads were assembled using PEAR (Zhang et al., 2014) version 0.9.1, with minimum and maximum paired read lengths (nt) of 470 and 550 bp. In addition, no ambiguous bases (N) were allowed. Reads were further processed using the USEARCH (Edgar, 2010) pipeline with a maximum expected error set at 0.5, and a minimum phred score of 20. *De novo* chimera detection was performed using UCHIME (Edgar et al., 2011) and reads were clustered into Operational Taxonomic Units (OTUs) at a 97 % similarity cut-off using UPARSE (Edgar, 2013). The most abundant sequence in each OTU was used as a representative to which a taxonomy was assigned with Mothur's (Schloss et al., 2009) implementation of the Wang Naïve Bayesian Classifier (Wang et al., 2007), using a local Greengenes trainingset (version of May 2013) (McDonald et al., 2012) for prokaryotes. The sequences in this training set were trimmed to the first 570 bp (i.e., so that every reference contained the complete V1-V3 region), in order to reduce both incorrect identification and the number of unclassified sequences (Newton and Roeselers, 2012; Werner et al., 2012).

Downstream analysis was performed using the R environment (R Core Team, 2015), with the packages *Vegan* 2.3-0 (Oksanen et al., 2015) for redundancy analysis (Hellinger-transformed data and scaled abiotic variables), non-metric multidimensional scaling (NMDS), calculation of PCNM variables (Borcard et al., 2004) and diversity and distance metrices, and *clustsig* 1.1 (Whitaker and Christman, 2014) for a SIMPROF analysis (Clarke et al., 2008) with 999 iterations, alpha set to 0.05 and average clustering using the Bray-Curtis dissimilarity.

4.4. Results

After pre-processing, about 3 million high quality reads distributed over 15,649 OTUs were kept. Before the downstream analysis, we further removed archaean, chloroplast and mitochondrial OTUs, strict single- and doubletons (i.e. OTUs represented by at most 2 sequences in the entire dataset), and those samples having less than 2073 sequences, leaving us a final 138 samples. Just over 2.56 million sequences and 5954 OTUs were retained, divided over 45 phyla and 325 genera. At the phylum level, 8807 sequences (0.3 %) in 323 OTUs (5.4 %) remained unclassified, while this was the case for 765,492 sequences (29.8 %) in 4397 OTUs (73.8 %) at the genus level. Samples had on average 18,570 ± 19,861 sequences and 287 ± 173 OTUs.

Cyanobacteria and *Proteobacteria* showed the highest relative abundances (Fig. 4.2), having at least 10 (*Proteobacteria*) to 20-fold (*Cyanobacteria*) more sequences than the next two groups (*Actinobacteria* and *Bacteroidetes*). A further decline in relative abundances was observed for the phyla *Firmicutes* and *Chloroflexi*, and *Acidobacteria* and *Planctomycetes*, while 38 additional phyla were less represented still. Most OTUs belonged to the *Proteobacteria* (31 %), while *Cyanobacteria*, *Actinobacteria* and *Firmicutes* were represented by relatively few OTUs considering their number of sequences (respectively 5.9 %, 7.2 % and 4.6 % of the total). In contrast, *Bacteroidetes*, *Chloroflexi*, *Acidobacteria*, *Planctomycetes*, *TM7*, Candidate phylum *OD1* and *Armatimonadetes* displayed a relatively high OTU-richness.

Although Continental and Maritime Antarctica showed similar relative abundances of ~60 % for *Cyanobacteria* and ~20-25 % for *Alphaproteobacteria*, *Firmicutes* were more abundant in the Maritime samples, while *Actinobacteria* were more important in the Continental samples (Figures S4.1 and S4.2). The potentially larger productivity in Macquarie Island by eukaryotic phototrophs could explain the higher abundance of heterotrophic *Alphaproteobacteria*, and

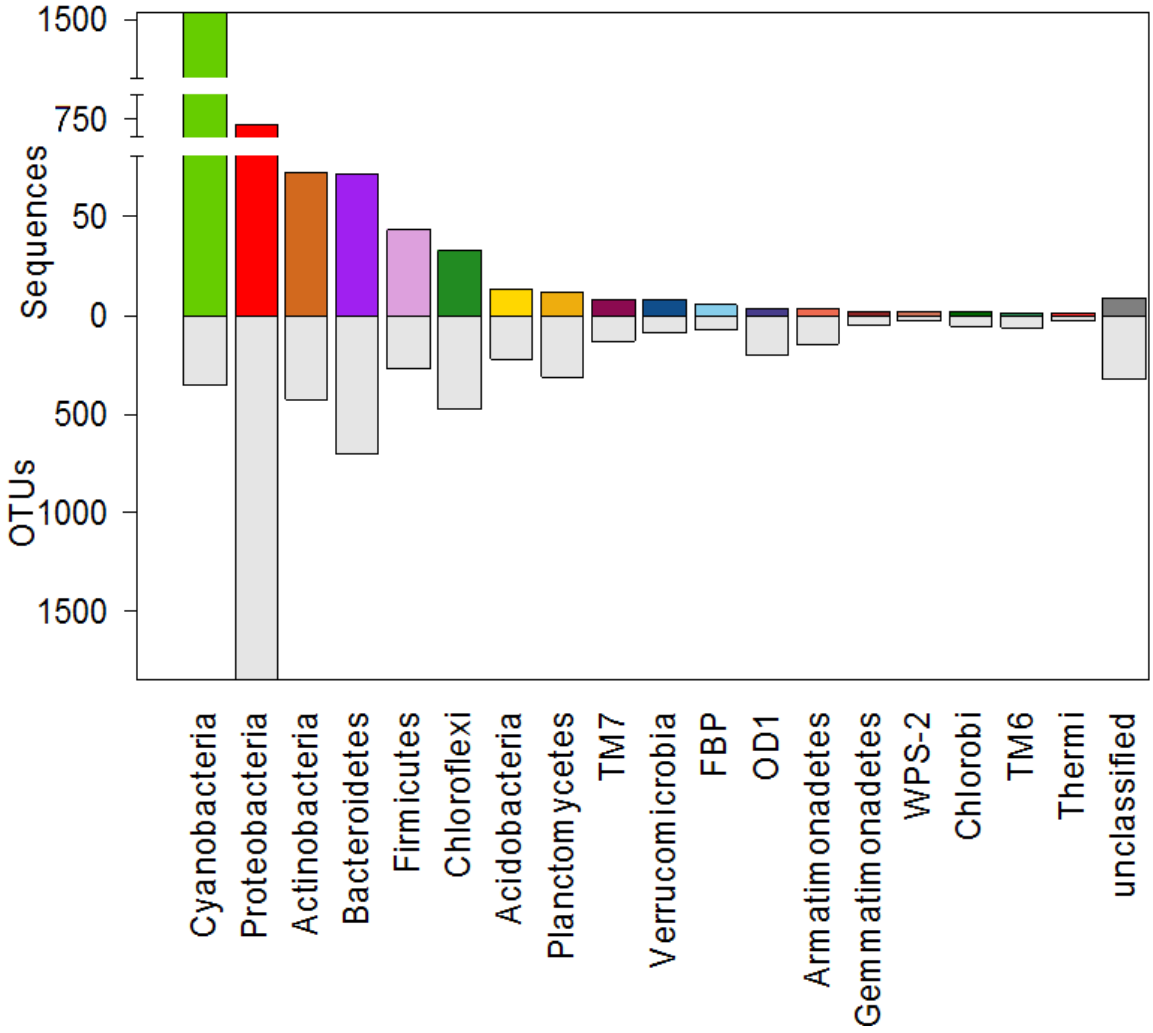

Figure 4.2. The number of reads and OTUs for the 18 best represented phyla (> 1000 sequences) and unidentified OTUs in the total dataset. Read numbers of *Cyanobacteria* and *Proteobacteria* should be multiplied by 1000, those of the other phyla by 100.

for example the presence of the *Koribacteraceae* family of phylum *Acidobacteria*, which were largely restricted to this island, and the relatively lower number of *Cyanobacteria*. In contrast, Marion Island samples were dominated by *Betaproteobacteria* and *Firmicutes* (60 % and 16 % of the reads, respectively), while *Cyanobacteria* made up only about 5 % of the reads.

A heatmap (Fig. 4.3) showing the Bray-Curtis dissimilarities (D_{BC}) revealed a generally low similarity between samples, with an average dissimilarity of 0.92 ± 0.14 for the relative

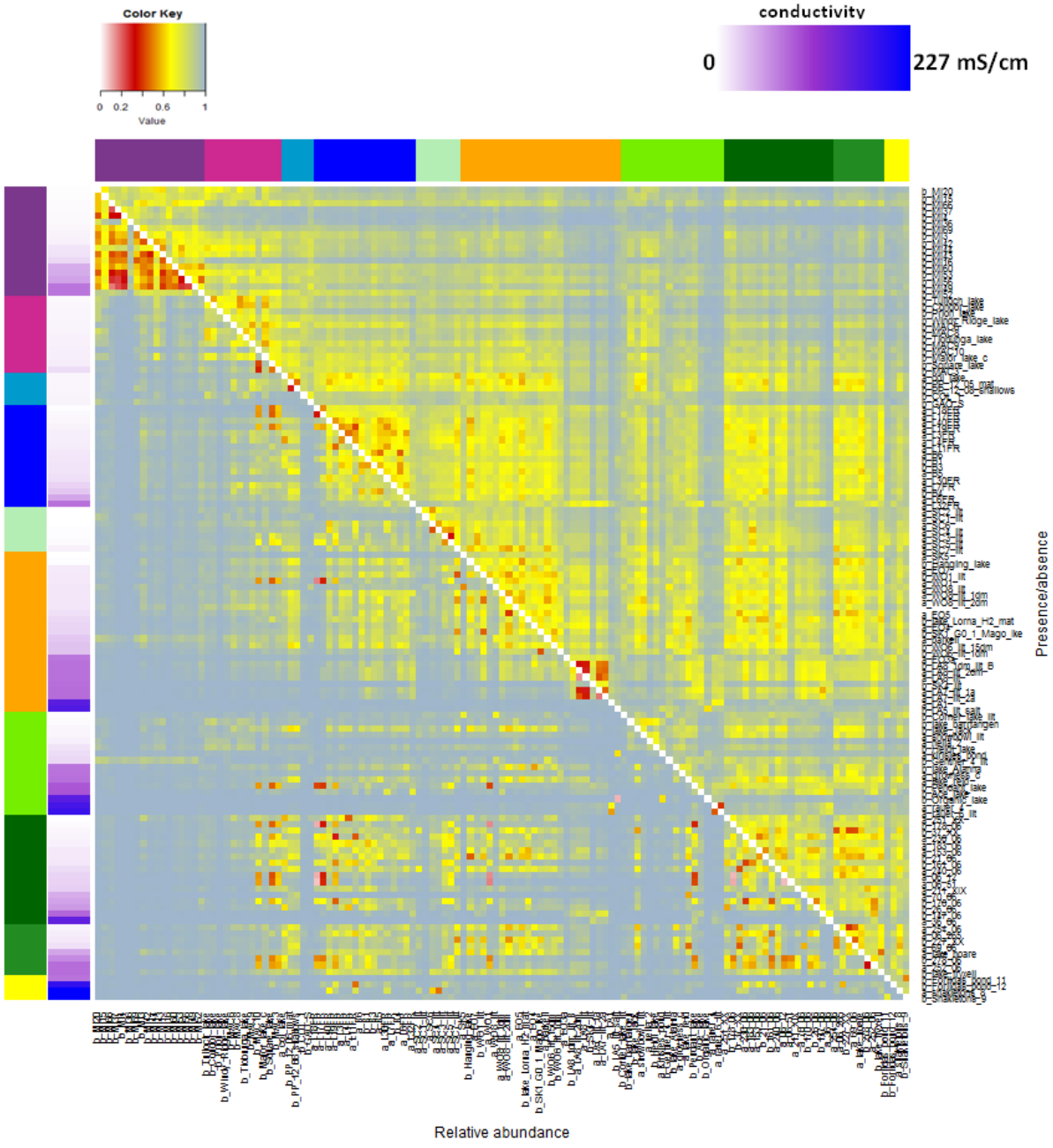

Figure 4.3. Heatmap showing the Bray-Curtis dissimilarity between samples. Samples are grouped by region (upper colour scale) and then ranked according to increasing electric conductivity within a region (purple-blue colour scale). Region order is Marion Island (purple), Macquarie Island (fuchsia), West-Peninsula (light blue), East Peninsula (dark blue), Dronning Maud Land (seagreen), Enderby Land (orange), East Antarctica (bright green), North Victoria Land (dark green), South Victoria Land (light green) and the Transantarctic Mountains (yellow). The lower triangle is based on relative abundances, the upper triangle based on presence/absence data.

abundance data, and a slightly lower dissimilarity for the presence-absence (PA) data ($\bar{D}_{BC}$ = 0.83 ± 0.13). When considering the relative abundance data, Marion Island showed a

high dissimilarity to other regions, although PA data revealed a moderate similarity to Macquarie Island. Overall, although similarity was relatively low, continental samples with a low to moderate EC tended to show a higher similarity to each other, while samples with a high EC, such as the Rauer lakes, showed nearly no similarity to most samples. Relatively higher similarities were furthermore observed between samples within an oasis with more or less similar EC values, such as the Langhovde samples in Enderby Land.

When using PA data, a higher overall similarity became evident, suggesting that many OTUs were relatively widespread albeit generally rare (mostly < 10 reads per sample when present). The PA data also pronounced the observed spatial autocorrelation, which was also suggested by the clustering and NMDS analyses.

Clustering and NMDS

An NMDS analysis (Figure 4.4a) showed no unequivocal geographical separation between the different regions (stress = 0.21), largely with the exception of Marion Island, which did display a more coherent clustering, and to a lesser extent also the Macquarie and TAM samples. Nevertheless, samples from the same region tended to cluster closer to each other, which was also corroborated by the SIMPROF-analysis, showing larger, significant clusters of Marion Island, Macquarie Island and Maritime samples (Fig. S4.3). The samples from East Antarctica seemed to be an exception, and displayed a more scattered distribution. A slightly more coherent geographical clustering became apparent when using presence-absence (PA) data (Figures 4.4b and S4.4). Also, high (“salt”) and moderately high (“brackish”) electric conductivity samples did generally not group with the freshwater samples. This was also more obvious from the cluster analysis, where samples with a high electric conductivity clustered separately as an outgroup.

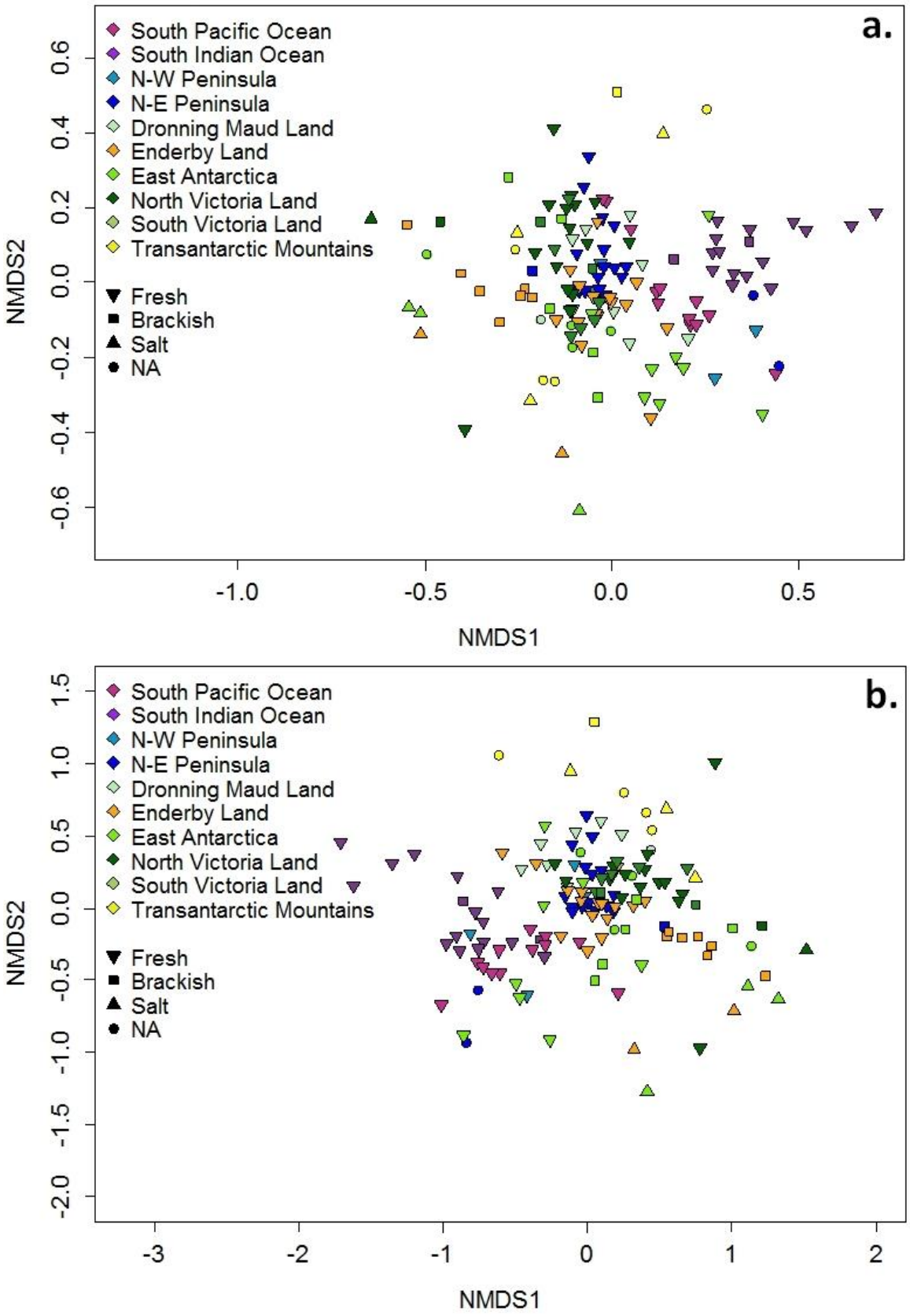

Figure. 4.4. NMDS analyses using the Bray-Curtis dissimilarity. The data was rarefied and transformed to **(a)** relative abundances or **(b)** presence-absence. Colours represent the region, and symbols represent electric conductivity values: fresh ≤ 1.5 mS/cm; brackish 1.5 ≤ 14 mS/cm; saline > 14 mS/cm. Stress = 0.21 (a); Stress =0.20 (b).

In order to explore the effects of the environment and geographical distance, we performed a redundancy analysis (RDA) based on the parameters which retained the most samples. Using electric conductivity, pH, altitude, distance from the sea and mean annual temperature, we were able to keep a representative set of 117 samples (no information on temperature for the TAM was available, hence completely excluding this region). EC, distance from the sea and altitude were log-transformed in order to normalize them, after which all parameters were scaled. Coordinates were transformed into principal coordinates of neighbour matrices (PCNM), creating variables that correspond to all the spatial scales that can be perceived in

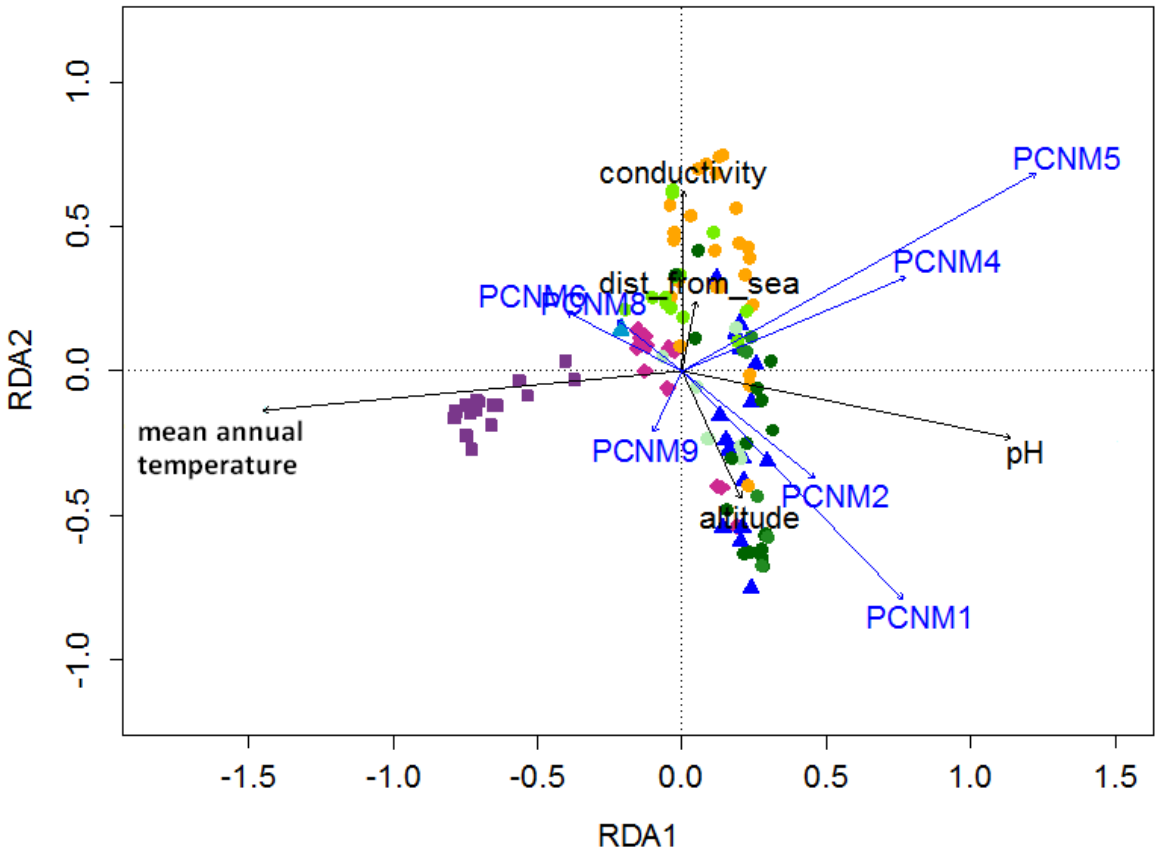

Figure 4.5. Redundancy analysis with 117 samples for which data was available on electric conductivity, pH, altitude, distance to the sea and average annual temperature. Only significant parameters are plotted. Region colours as mentioned earlier.

the data set (Borcard et al., 2004). Ordination was performed on Hellinger-transformed sequencing data. All parameters were highly significant ($P \leq 0.01$), except altitude, which was only marginally significant ($P = 0.03$). Of the 10 PCNM variables, 6 proved highly significant ($P < 0.01$), while PCNM8 was marginally significant ($P = 0.025$). Our model explained 23 % of the observed variation (R^2_{adj}). Variation partitioning further showed that spatial distance by itself explained 9 % of the variation, the environmental parameters 4% and the overlap between both 10 %. Mean annual temperature was the most important environmental parameter, followed by conductivity, pH, altitude and distance from the sea. The first RDA axis largely coincided with mean annual temperature and to a lesser extent pH, while the second axis nearly completely coincided with conductivity (Fig. 4.5). The PCNM variables explaining most variation were those representing medium to large geographical distances (PCNMs 5, 4, 1 and 2).

The RDA was repeated with either Marion Island removed (Fig. S4.5a) or both Sub-Antarctic islands removed (Fig. S4.5b). Temperature was still the most important variable when Macquarie Island was included, while the distance from the sea was no longer significant. When removing both Sub-Antarctic islands, neither distance from the sea, nor altitude were significant, while pH apparently was the most important variable.

Latitudinal richness gradient

A wide range in the OTU-richness could be observed within a certain region or at any specific latitude (Fig. 4.6). The numbers of OTUs in the TAM samples were on average reduced compared to those from lower latitudes, although the richness in more northerly samples could be lower (e.g., North Victoria sample 70_06, of which 99 % of the sequences belonged to a single *Leptolyngbya* (*Cyanobacteria*, *Pseudoanabaenales*) OTU). Despite being the most northerly site (~46.93° S), samples from Marion Island showed a generally low

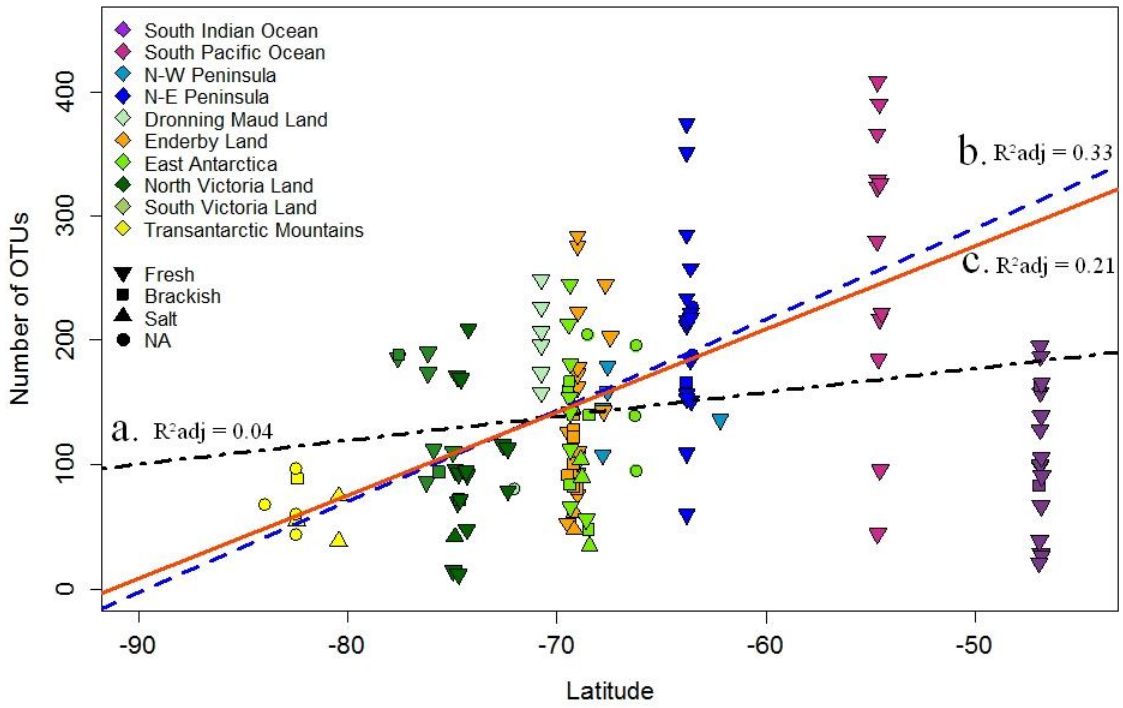

Figure 4.6. The number of OTUs of per sample (rarefied to 2073 sequences/sample) plotted against their latitudinal position. Linear regression showed a significant inverse correlation of OTU-richness with latitude: a. all regions (black dot-dash line); b. after the removal of Marion Island (blue dashed line); c. only Antarctica (full orange-red line). ($P < 0.01$ for (a); $P < 0.001$ for (b) and (c)).

richness with on average 154 ± 84 OTUs (ranging from 25 to 287, non-rarefied data). Contrary to the majority of the samples from other regions, which contained predominantly *Cyanobacteria* or *Alphaproteobacteria* (Fig. S4.2), most MI samples were dominated by *Betaproteobacteria*, in particular by OTUs related to *Janthinobacterium lividum* and congenerics (100 % identity with *Janthinobacterium* sp. TP-Snow-C76, Acc. KC987006 (BLAST NCBI)), and to a lesser extent *Firmicutes* (genus *Clostridium* or related *Clostridiaceae*). OTU-richness only showed a low inverse correlation with latitude when all samples were incorporated ($R^2_{adj} = 0.04$; $P < 0.01$) (Fig. 4.5, line a). However, when excluding the Marion Island samples from the regression analysis, a moderate but highly significant increase in OTU-richness with decreasing latitude was noticed ($R^2_{adj} = 0.33$; $P <$

0.001; line b), a trend which did not change when removing both Sub-Antarctic Islands (R^2_{adj} = 0.21; $P < 0.001$; line c) . To eliminate the possible effects of electric conductivity on OTU-richness, we excluded the extreme conductivity values (at both low and high ends), retaining only freshwater samples with an EC between 0.15 and 1.5 mS/cm (Fig. S4.6). This showed that the trend was consistent.

4.5. Discussion

This first large scale study using a uniform and high-throughput approach covering benthic microbial mat samples from lakes in all eight ACBRs (Terauds et al., 2012) where bedrock-based surface lakes have been found (Fig. 4.1; Table 4.1) and two Sub-Antarctic islands revealed that *Cyanobacteria* was the best represented phylum in most of our samples (Fig. S4.2). However, this phylum was less prominent in Marion Island where lakes were dominated by *Betaproteobacteria*. By contrast, *Cyanobacteria* showed very low relative abundances and richness in most high conductivity lakes, and generally a low richness in brackish lakes, except in the Antarctic Peninsula (Fig. S4.2). This is in agreement with other studies from polar lakes which revealed that *Cyanobacteria* typically dominate microbial mats and account for a large proportion in their biomass and productivity (de los Ríos et al., 2015). Cyanobacterial diversity is usually studied through a range of microscopy techniques (Strunecký et al., 2012; de los Ríos et al., 2015; Sumner et al., 2015). Molecular approaches, however, have revealed a higher diversity, and also suggest a higher level of endemism than previously assumed based on morphotypes (Taton et al., 2003; Strunecký et al., 2012). A recent analysis (September 2015) of all available cyanobacterial sequences in the NCBI genetic sequence database (GenBank) has resulted in 424 OTUs at a 97 % similarity cut-off in Antarctica (Zorigto Namsaraev, pers. comm.). Our results, too, suggest a relatively low

diversity with only 351 OTUs found for over 1.54 million sequences. This similar amount of OTUs between GenBank and our study also supports our data processing approach. A *Phormidium*-OTU was recovered from 109 samples and was present in every region, and had the overall highest read number of any OTU, yet was largely lacking in the Marion Island samples. *Leptolyngbya* OTUs were among the most widespread and abundant OTUs (up to 83 samples per OTU), and are known to typically dominate benthic microbial mats (Adams et al., 2006; Sumner et al., 2015). To a lesser extent, some *Pseudanabaena* and a *Nodularia sphaerocarpa*-related OTUs were found in up to 73 samples. Interestingly, an unclassified *Cyanobacteria* OTU recovered from 74 samples was the third most abundant OTU (6 % of total reads), and showed a 100 % identity (95 % length) with a recent clone from Lake Vanda (MDV) benthic microbial mats (Zhang et al., 2015) and 99 % similarity to a Lake Fryxell clone (Taton et al., 2003).

Proteobacteria were the second most abundant group, with especially the *Alphaproteobacteria* being numerically dominant. This is not surprising, given the high availability of organic matter in benthic microbial mat or biofilm systems. With 1114 OTUs, *Proteobacteria* were also the most diverse group. The overall most widespread OTU was a *Zymomonas* sp., recovered from 118 samples. *Alphaproteobacteria* OTUs in general were widely distributed, especially OTUs associated with the genera *Rhodobacter*, *Mycoplana*, *Erythromicrobium*, *Devosia*, *Kaistobacter*, *Roseococcus*, *Mesorhizobium* and *Bradyrhizobium*. Many of these are (facultative) anaerobic and some could possibly be involved in (anoxygenic) photosynthesis (*Rhodobacter*, *Roseococcus*, *Erythromicrobium*) or nitrogen-fixing (*Mesorhizobium*, *Bradyrhizobium*, *Devosia*), while others are aerobic heterotrophs, such as *Mycoplana*.

The recently proposed candidate phylum *FBP* (previously OP11), closely related to phylum *Armatimonadetes* (Lee et al., 2013), was nearly completely restricted to the Antarctic

mainland, with the exception of four OTUs recovered from Macquarie Island in very low abundances (4 of 5 instances were singletons). In total, this phylum was represented by 69 OTUs, and a relatively high richness and abundance was present in the eastern Peninsula and Dronning Maud samples. Blasting representative sequences revealed that this phylum appears to be mainly recovered from polar and alpine regions, although some matches came from more temperate regions as well. They may, however, be primarily a terrestrial phylum, as their richness recovered from terrestrial samples (Chapter 3) is relatively much higher (90 OTUs from 52 samples within one region).

Biogeographical and macroecological patterns

Our region wide study revealed an on average increasing number of OTUs in each lake between 85 and 54° S, while the richness decreased again in the water bodies in Marion Island (46° S). With the exception of Marion Island, the patterns observed are in agreement with one of the most generally observed patterns in ecology (Willig et al., 2003) and were also reported in Antarctic macrophytes (Peat et al., 2007) and macrofauna. Conclusive data on the existence of a latitudinal diversity gradient in microorganisms has been generally lacking (Lawley et al., 2004; Yergeau, Newsham, et al., 2007; Fontaneto et al., 2015), but see Vyverman et al. (2007). Because of their huge abundances, minute sizes and hence traditionally accepted high dispersal capacities, it was hypothesized that microorganisms would display no or only a weak latitudinal diversity gradient, with many cosmopolitan species also found at higher latitudes (Vincent, 2000; Fenchel and Finlay, 2004; Tindall, 2004). Nonetheless, evidence for a latitudinal diversity gradient has been demonstrated for bacterioplankton (Schiaffino et al., 2011) and bacteria in bare soil (Yergeau, Newsham, et al., 2007) as well as lacustrine diatom communities (Vyverman et al. 2007). However, exceptions to this general pattern have been reported in for example Mars Oasis, which is considered to be a biodiversity hotspot. The trend was, however, invalid when plant cover was present

(Yergeau, Newsham, et al., 2007). Despite the decrease in OTU richness with latitude, a lot of variation was visible at any particular latitude or within any region (Fig. 4.6), probably as a result of local (micro-scale) environmental conditions. This was also observed in a study in Victoria Land where local conditions were shown to complicate interpretations and could overwhelm possible latitudinal trends (Howard-Williams et al., 2010).

At the highest end of the latitudinal gradient, the TAM lakes displayed the overall lowest OTU richness of any region (66 ± 19 OTUs, rarefied). Several studies have reported generally depauperate invertebrate, cyanobacterial and cryptogamic richness in the TAM compared to coastal sites (Hodgson et al., 2010; Fernandez-Carazo et al., 2011; Magalhães et al., 2012). This can be assigned to both the geographical isolation and multiple environmental stressors (Hodgson et al., 2010), such as extremely high UV-radiation and low temperatures. It is important to note that freshwater samples were not available from this region, with lakes showing moderate (2.22 mS/cm)[5] to extremely high EC values (227 mS/cm), which likely additionally contributes to a lower richness (Fig. S4.8). Despite their geographical isolation, several OTUs from the TAM have previously been reported from other regions, such as cyanobacteria recovered from the East Antarctic saline Rauer lakes (Hodgson et al., 2010). It is evident from our data too, that many OTUs in these samples are widespread across the continent.

The highest average richness was found in the samples from Sub-Antarctic Macquarie Island (266 ± 115 OTUs, rarefied). Particularly, *Alphaproteobacteria* showed to be both very abundant and very diverse on the island, with 49 % of all alphaproteobacterial OTUs recovered from the merely 12 samples. Overall, nearly 46 % of all 1847 proteobacterial OTUs

[5] Littoral EC from the moat around Forlidas Pond, while the benthic zone had an EC of 142 mS/cm (Hodgson et al., 2010)

were retrieved from these samples. *Cyanobacteria* showed a relatively high richness (153 OTUs), but were numerically less important compared to the mainland samples.

Marion Island proved to be an anomaly. Despite being the northernmost site, having the highest precipitation rates and annual average temperatures (Le Roux and McGeoch, 2008), it showed a low richness (105 ± 57 OTUs, rarefied) compared to Macquarie Island, and even to the continental regions. Strikingly, the samples were dominated by few *Janthinobacterium*-related (*Betaproteobacteria*, *Oxalobacteraceae*) OTUs. These organisms are known to produce several antifungal, antiviral and antibacterial substances (e.g., violacein). By forming a biofilm they can suppress other organisms and enhance their own survival (Alonso-Sáez et al., 2014). However, there was no significant correlation between OTU-richness and the relative abundance of *Janthinobacterium* (data not shown). Interestingly, the most abundant OTU – identified as *Janthinobacterium lividum* – was recovered from 55 samples (all regions), although 97.3 % of the reads were confined to Marion Island. Three less numerous OTUs were limited to Marion Island, and a fifth *Janthinobacterium*-OTU was restricted to mainland Antarctica (i.e., peninsular and continental Antarctica). Several other *Oxalobacteriaceae*-OTUs were abundant in the MI samples (a BLAST search revealed a close affinity to *Duganella*). *Firmicutes* too were predominantly recovered from the MI samples, with OTUs related to *Clostridium* being the most abundant. These main *Firmicutes* OTUs were also recovered in relatively high numbers from the east Peninsula samples. In contrast, although 76 cyanobacterial OTUs were recovered (22 %), they were mostly present in very low abundances, with only 23 OTUs having more than 10 sequences. The relatively high abundance of *Firmicutes* and *Oxalobacteraceae*–related OTUs might indicate anaerobic conditions, although *Janthinobacterium* and *Duganella* are obligate aerobic bacteria (Hiraishi et al., 2015). Unfortunately, no data on oxygen-levels were available for the Marion Island samples. It is, however, unclear what could explain the generally low richness in these

samples and the dominance of *Janthinobacterium*. For example, high viral loads could temporarily affect bacteria with less efficient defence mechanisms, although it is likely to be a combination of factors.

In addition to the latitudinal diversity gradient, we also observed clear differences in the overall community structures between the Antarctic continent and surrounding Maritime Antarctic islands and the Sub-Antarctic islands (Fig. S4.5), but also within mainland Antarctica. Clustering of the samples (Figures S4.3 and S4.4) did not result in clearly delineated regions such as those proposed for macroorganisms (Terauds et al., 2012). Marion Island samples, and to a lesser extent the Macquarie and Maritime Antarctica samples showed a stronger clustering, while no clear differentiation for the Continental samples was apparent. Electric conductivity did seem to play a role, as the samples with very high EC clustered separately, while also samples in close proximity with similar EC values tended to cluster more closely. Furthermore, the comparison of the number of shared OTUs between the three biogeographic regions (Fig. S4.7) showed that Maritime and Continental Antarctica shared more OTUs than either did with the Sub-Antarctic. This could indicate that the proposed Gressitt Line between the Peninsula and the Continent (Chown and Convey, 2007), the equivalent of the Wallace Line in south-east Asia, might be not or less valid for, at least, benthic bacteria. In a meta-analysis of terrestrial studies, Chong et al. (2012) concluded that the differences in communities between Maritime/Scotia Arc and Victoria Land indeed validate this ancient biogeographic border for bacteria. However, our results indicate that while there are differences in communities between the Maritime Antarctic and e.g. Victoria Land samples, these differences are not necessarily larger than between Continental samples from different regions. Moreover, samples from different regions, including the Maritime Antarctic and Victoria Land, were shown to cluster together (Figures S4.3 and S4.4). This could indicate that indeed bacteria are less dispersal limited, although physical and

environmental barriers might prevent complete mixture. However, we should be aware that the resolution of our amplicon might not be sufficient, and that our OTU clustering is a distance-based approach. To completely resolve this we would need to perform phylogenetic analyses, preferably on longer fragments and using a multi-locus approach. Furthermore, evolutionary rates are likely to be reduced in polar taxa both directly because of the low temperature, and also because of longer generation times (Convey et al., 2014), so that differentiation between populations will be slower. Biogeographical zoning might hence in fact be higher still, even within the continent, although this would also require more information on the connectivity between the regions. It seems also likely that in e.g. meromictic lakes little or no exchange of organisms adapted to the conditions in the deepest zones will occur between lakes, leading to differentiation even within regions.

Much less OTUs were shared between Continental Antarctica and both islands, while this was less pronounced for the Peninsula, particularly sharing OTUs with Macquarie Island. Still, the number of OTUs shared between both Sub-Antarctic islands was higher. This might be due to both effects of isolation or connectivity between and within the Antarctic and Sub-Antarctic (Chong et al., 2015), as well as environmental differences, particularly temperature and moisture availability and, related to this, the presence of catchment vegetation and higher dissolved organic carbon concentrations. The effect of temperature is for example obvious from the redundancy analysis (Fig. S4.5), which revealed a separation of the Sub-Antarctic islands and the western Peninsula from most continental samples. In addition, local environmental factors are important in structuring bacterial communities and restricting their composition (Chong et al., 2015). The coincidence of electric conductivity with the second RDA axis corroborated the observations of our similarity analyses, namely that EC appeared to be a major driver in structuring the communities, particularly within the continent. Lakes with similar EC showed a higher resemblance, generally independent of their geographical

distance, although, lakes in close proximity with comparable EC were more similar to each other. The observation that high conductivity lakes share similar bacterial assemblages has been made previously (Wilkins et al., 2013), with members of the phylum *Bacteroidetes* and class *Gammaproteobacteria* dominating lakes from both the Vestfold Hills (VFH) (Bowman et al., 2000) and MDV (Mosier et al., 2007), and with lower abundances of *Alphaproteobacteria*, *Firmicutes* and *Actinobacteria*. Surface waters of such lakes resemble marine communities, although with higher abundances of *Actinobacteria* and certain cyanobacterial groups (Lauro et al., 2011). *Bacteroidetes* were indeed more abundant and showed a higher OTU richness in the high EC samples (data not shown). Nevertheless, an indicator species analysis based on our conductivity classes showed a high number of *Rhodobacteraceae* (*Alphaproteobacteria*) OTUs to be indicative for both saltwater and brackish samples. The family *Rhodobacteraceae* are fundamentally aquatic bacteria that frequently thrive in marine environments (Pujalte et al., 2014), which might explain our observations.

In a metagenomics study, Yau et al. (2013) have found *Marinobacter*, *Roseovarius* and *Psychroflexus* to be the dominating bacteria in Organic Lake (VFH), a marine-derived hypersaline lake that has the highest reported dimethylsulfide (DMS) concentration in a natural water body. Our query showed a low richness (31 OTUs), dominated by a *Ralstonia* sp. (77 % of reads), which was also found in other high EC samples. *Marinobacter* was not recovered from Organic Lake, but was found in both *Rauer* lakes, Langhovde 1 (Enderby Land) and Bunger Hills lake 16, all high-salinity lakes. Likewise, *Psychroflexus* was recovered from either the same high-salinity lakes, or freshwater lakes in close proximity to the sea. The presence of this genus in the latter lakes might therefore result from sea spray. *Roseovarius* was not identified in our dataset, although many other *Rhodobacteraceae* were. In summary, it is becoming evident that for bacteria, distance between sites does effect

community compositions and that biogeographic zoning is at play. However, environmental factors might locally be more important, with especially electric conductivity appearing to be a strong structuring force. Despite a distance effect, lakes with similar limnological conditions from different regions did show a higher similarity. This was also observed in bacterioplankton over large geographical gradients (Van der Gucht et al., 2007), but for example not in diatoms or other micro-algae (Vyverman et al., 2007; Verleyen et al., 2009, unpubl.res.; Sassenhagen et al., 2015).

Microorganisms are thought not to be dispersal limited by their size *per se* (Martiny et al., 2006), although several effects might play between different groups of organisms. Long distance atmospheric dispersal is expected to be enhanced when the organisms are able to cope with the cold and dry conditions during the transport. For example, spore-forming or Gram-positive bacteria will likely to have higher survival rates, as was shown for e.g. *Firmicutes* and *Actinobacteria* (Smith et al., 2013). Contrary, when organisms are only capable of (active) short distance dispersal, it might take many generations to spread great distances, leading to genetically divergent populations (Martiny et al., 2006). Niche or habitat preferences will also play a role, as subsurface organisms (or benthic) are less likely to be dispersed, while habitat availability (partly) determines success of establishment. Organisms of extreme and rare habitats will hence likely show more endemism.

4.6. Conclusions

It is now been recognized that Antarctica's biodiversity is much more extensive, ecologically diverse and biogeographically structured than previously thought. Indeed, our survey seems to corroborate this observation, with a biogeographic zoning between the Sub-Antarctic islands and Antarctica being present, possibly related to differences in catchment characteristics and

mean annual temperature. Turnover in bacterial communities on the continent are to a large extent related to specific conductance, additional to geographic connectivity. A high overall diversity, with many OTUs unclassified and potentially new to science was observed. *Cyanobacteria* and *Alphaproteobacteria* appeared to dominate most freshwater samples, while *Bacteroidetes* and *Rhodobacteraceae* (*Alphaproteobacteria*) were more important in high-conductivity lakes. Marion Island was dominated by *Janthinobacterium* (*Betaprotobacteria*) related OTUs. OTU richness showed an inverse correlation with latitude, although a lot of variation within a region was observed, probably due to micro-scale environmental or limnological differences.

4.7. **Supplementary figures**

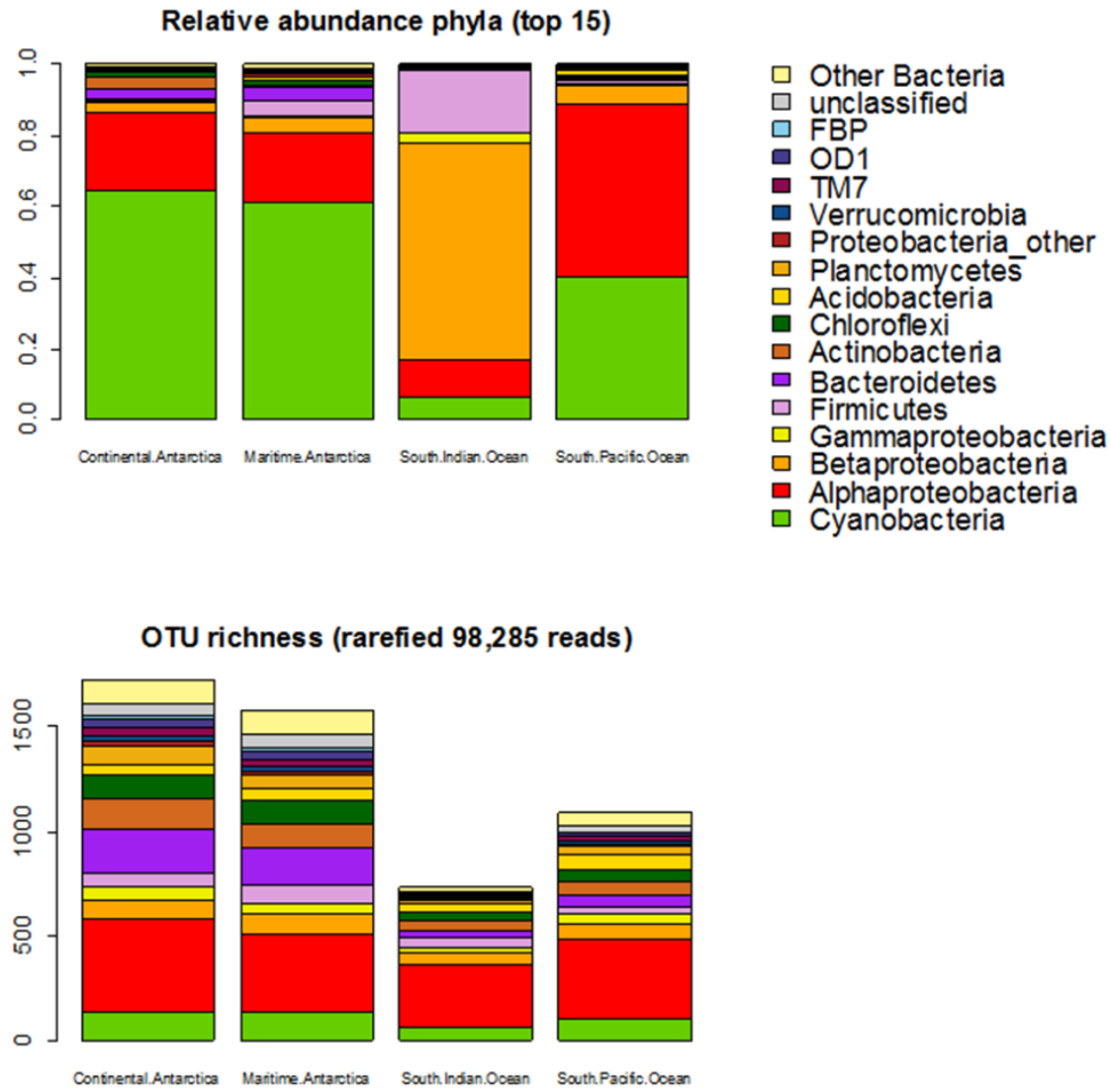

Figure S4.1. Summary of (a) the relative abundance and (b) the number of OTUs of the 15 most abundant phyla, the unclassified OTUs at phylum level and the sum of the remaining phyla for the four main regions (Continental Antarctica, the Antarctic Peninsula, South Indian Ocean (Marion Island) and the South Pacific Ocean (Macquarie Island)).

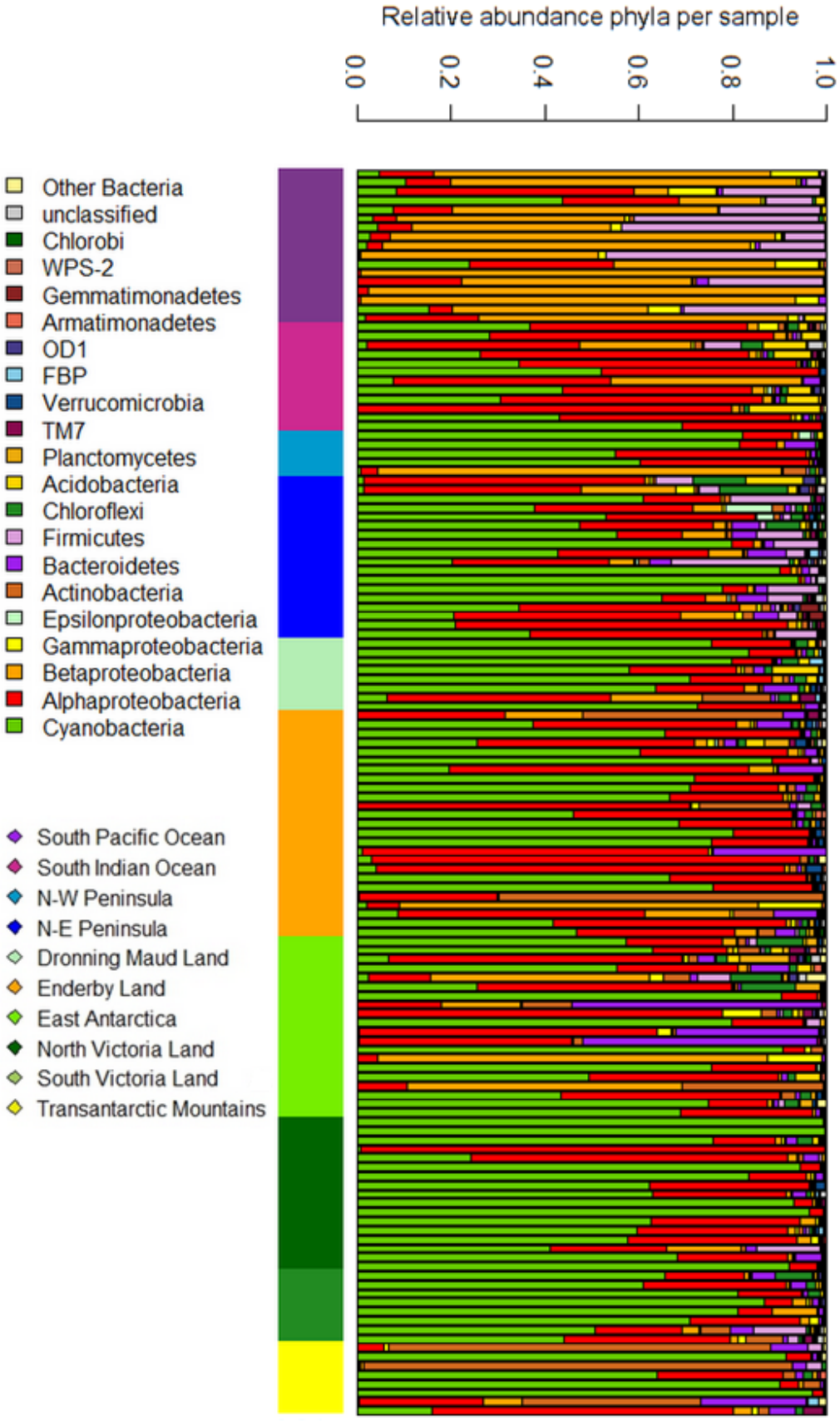

Figure S4.2. Relative abundances of the phyla per sample, grouped per region.

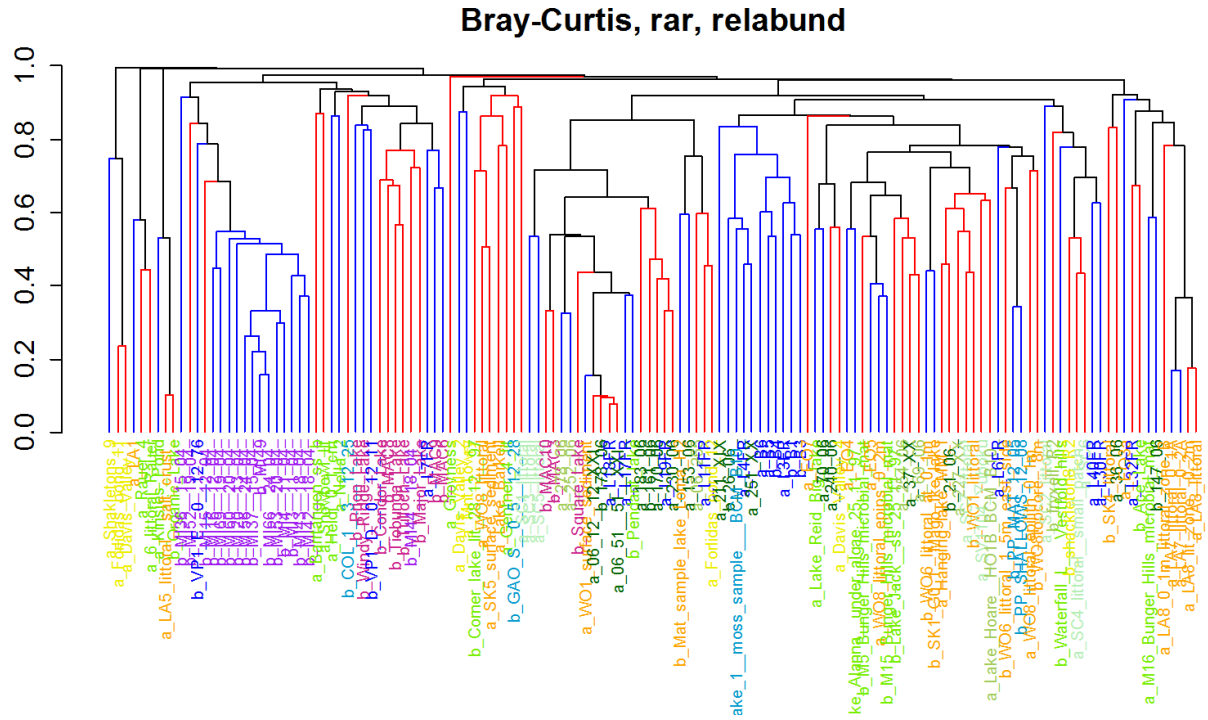

Figure S4.3. Average clustering of the relative abundances of rarefied data using the Bray-Curtis distance. **SIMPROF analysis** was used to reveal significant clusters (black lines). Colour of the regions as mentioned higher.

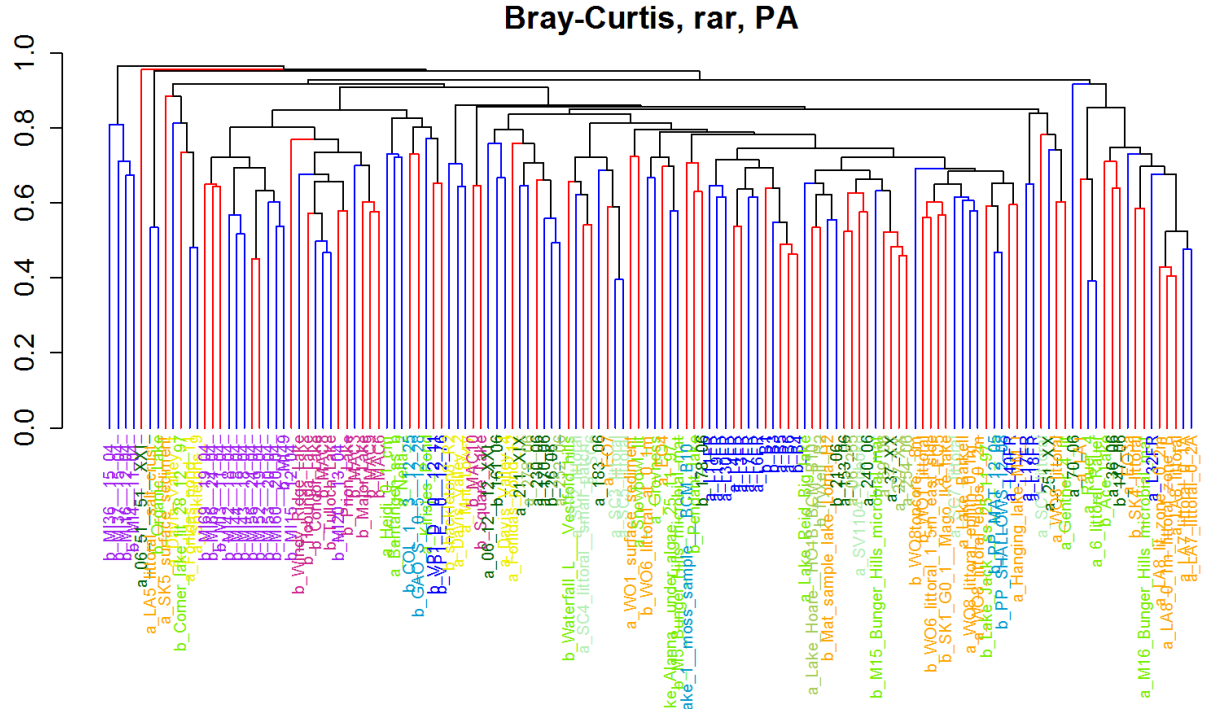

Figure S4.4. Average clustering of the rarefied data using the Bray-Curtis distance calculated on presence-absence values. **SIMPROF analysis** was used to reveal significant clusters (black lines). Colour of the regions as mentioned higher.

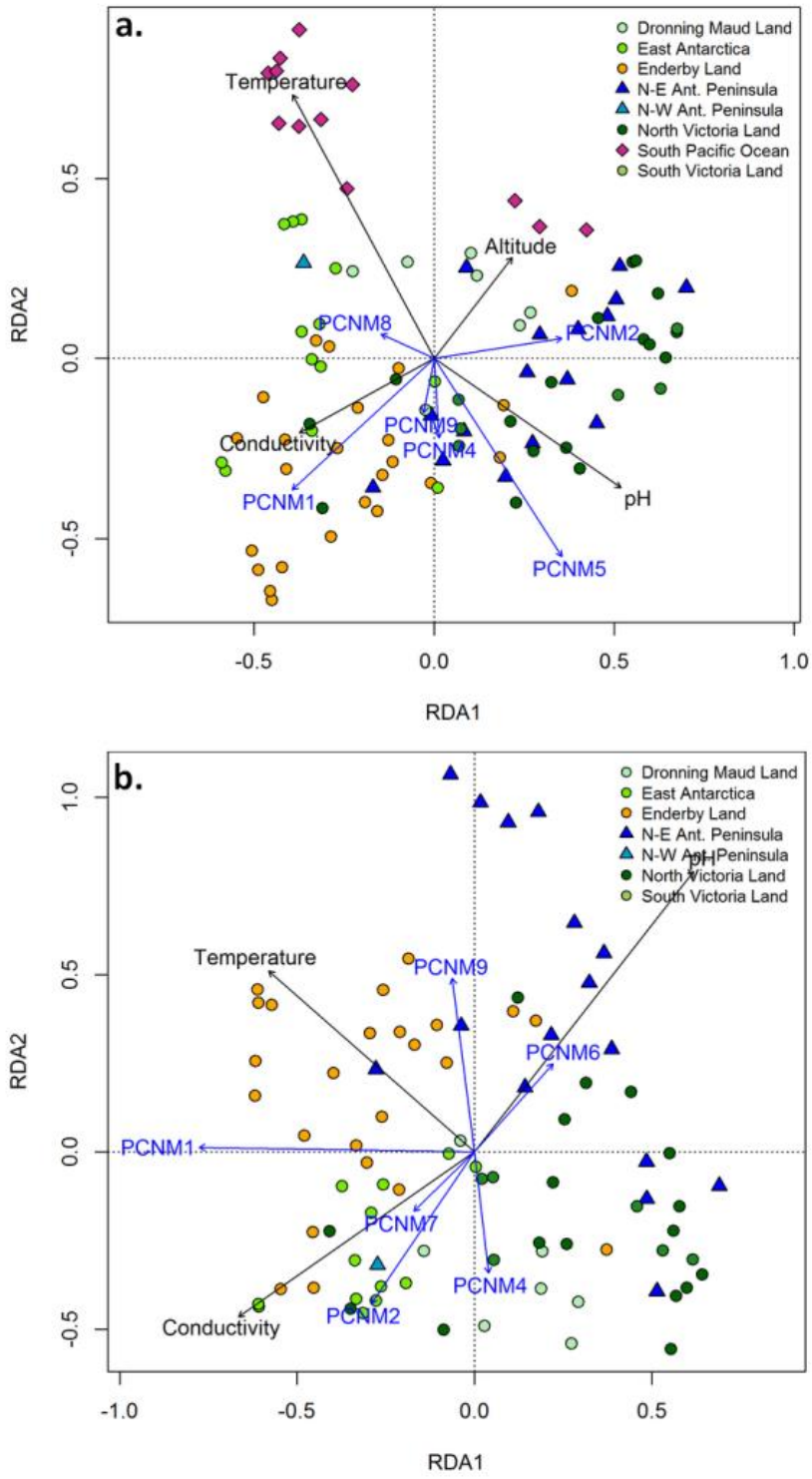

Figure S4.5. Additional RDA analyses (a) without Marion Island; (b) without Marion Island and Macquarie Island.

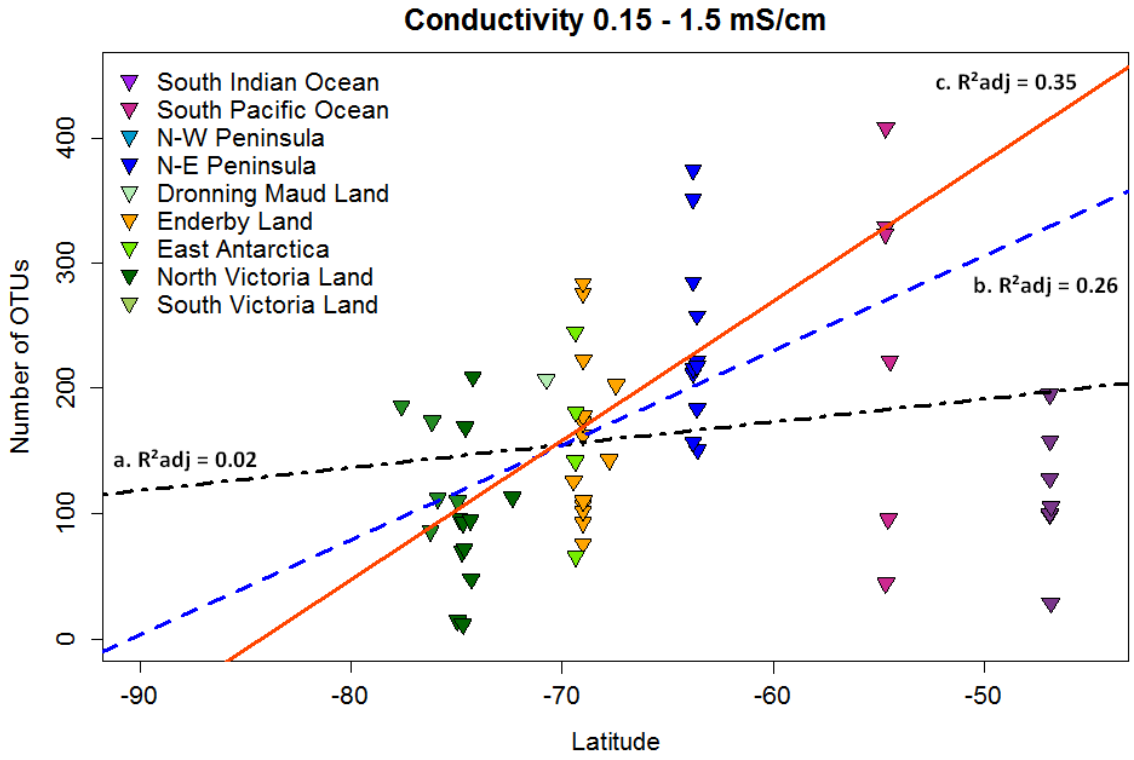

Figure S4.6. Number of OTUs in function of latitude of the freshwater samples. The samples with extremely low EC (< 0.15 mS/cm) were removed. The latitudinal pattern remains visible. a. all regions (black dot-dash line); b. after the removal of Marion Island (blue dashed line); c. only Antarctica (full orange-red line). All *P*-values significant (not displayed).

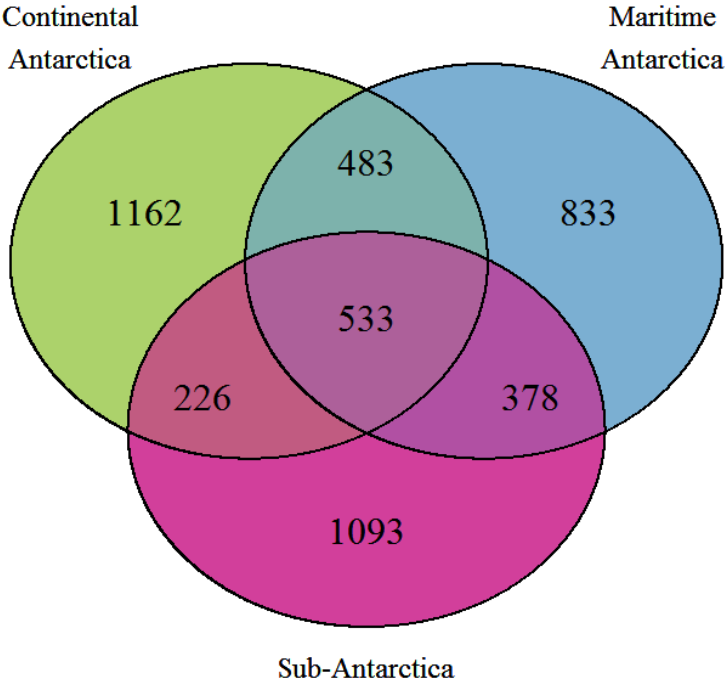

Figure S4.7. Venn diagram showing the number of shared OTUs between the biogeographical regions, and those unique for a region (rarefied to 96,394 reads per region).

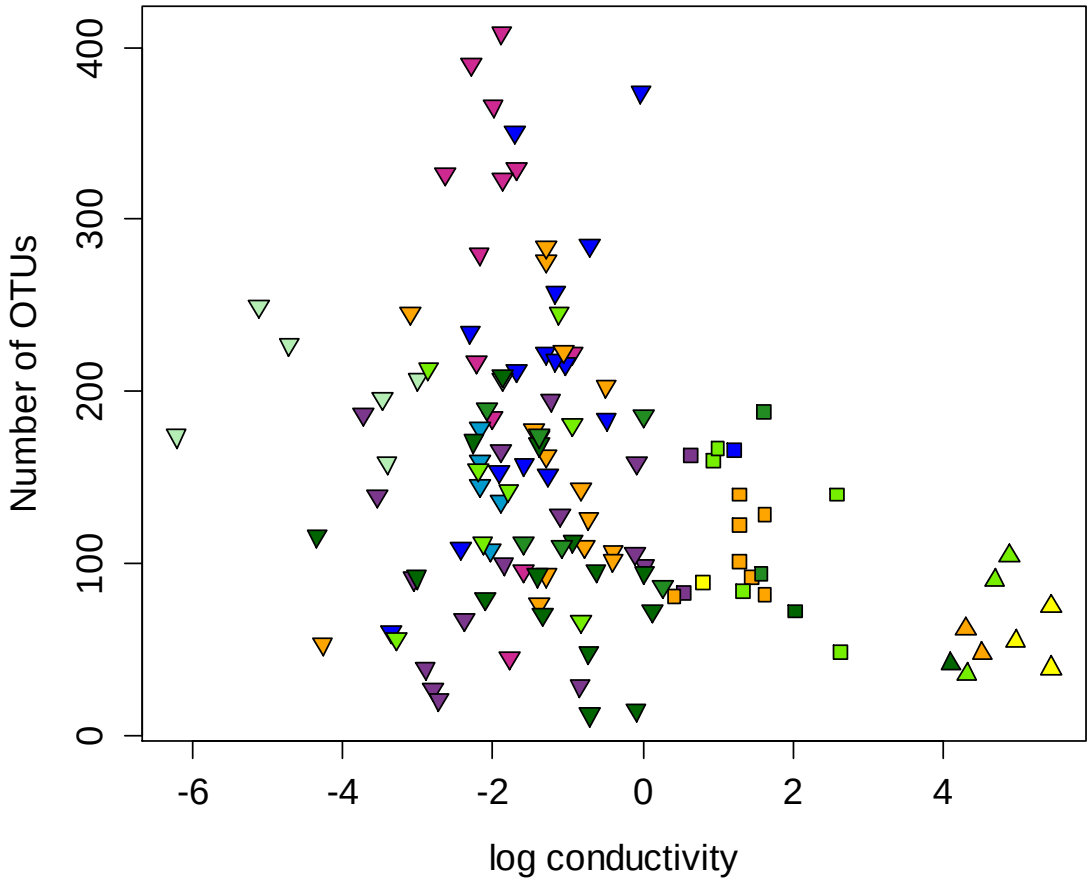

Figure S4.8. Number of OTUs in function of the log-transformed conductivity. A unimodal distribution is noticeable, with especially a trend towards reduction in OTU richness with higher EC. (R^2_{adj} = 0.11; P < 0.01).

Chapter 5. General discussion and perspectives

The advent of high-throughput sequencing (HTS) approaches, allowing a much higher resolution than more traditional culture-independent techniques at a continuously decreasing per-nucleotide cost, has enabled to study microbial communities with an unprecedented detail (Prosser, 2015). Based on pioneering studies a much larger bacterial diversity was suggested than that based on traditional methods (e.g., Sogin et al., 2006). However later it was shown that much of this unknown diversity in these early studies spawned from sequencing errors and that these approaches need a rigorous data processing and quality control (Huse et al., 2010; Kunin et al., 2010). Very recently, a lot of effort has been put into resolving the bacterial biodiversity in various environments, including the human gut (Huttenhower et al., 2012), the Arctic Ocean (Kirchman et al., 2010), petroleum contaminated soils (Bell et al., 2011), sulfur springs (Youssef et al., 2012), debris-covered alpine glaciers (Franzetti et al., 2013) and the ant microbiome (Kautz et al., 2013). By contrast, some ecosystems which are known to be dominated by microorganisms remain largely understudied. This for example the case with lacustrine and terrestrial habitats in Antarctica. This is surprising given the exceptional location and extreme environmental conditions which have led to an extraordinary adapted, although highly impoverished biodiversity which is highly dominated by microorganisms from all domains of life (Convey et al. 2015). Yet, exactly the relative simplicity of these ecosystems could allow biologists to reach a deep understanding of their functioning, and use them as models from which to extrapolate to more complex systems. Because of this relative simplicity, the functional roles of Antarctic microorganisms in general, and bacteria in particular, are especially significant (Wynn-Williams, 1996; Chong et al., 2015). The study of Antarctic bacteria is hence interesting for several reasons. First, the adaptation potential, physiology and functioning under the extreme environmental conditions

is largely unknown and might be interesting from a biotechnological point of view Second, these relatively simple systems allow to model the interactions between different bacterial groups and with other domains of life. Third, their functional, taxonomic and phylogenetic diversity; and fourth, their origin: to what extent are they pre-glacial remnants or recent colonists, and does their distribution provide clues? However, while relatively little research has been performed on Antarctic microorganisms compared to other regions, and little is known about their diversity and distribution, even only 6 % of all studies on Antarctic microorganisms deal with bacteria and archaeans, and few have actually been based on the 16S rRNA gene (Chong et al., 2015). It is evident that much is to be discovered in the Antarctic. Studies have mainly been limited to a few regions, generally near the continent's fringes where research stations are located. Many (inland) lakes and soils remain under sampled, while other habitats such as epiglacial systems, ice and snow have barely been investigated at all. Since over 99 % of the continent is estimated to be covered with ice, this vast surface and volume and the unexplored subglacial systems might result in unique discoveries. Additionally, the use of a range of techniques and approaches make a general comparison between regions and studies impossible.

In this PhD study, two HTS platforms (Roche's 454 pyrosequencing and Illumina's MiSeq) have been employed in order to (i) increase the general knowledge on bacterial diversity in Antarctica, (ii) investigate the bacterial community composition of soils in the Sør Rondane Mountains and microbial mats from both the Antarctic and Sub-Antarctic, and link this to the available environmental parameters, and (iii), assess biogeographical patterns. The results obtained in HTS were compared to either previous heterotrophic cultivation approaches (Chapter 2) or a genetic fingerprinting technique (ARISA) (Chapter 3).

5.1. Bacterial diversity surveys benefit from complementary methods

While the actual diversity study is currently in preparation (Obbels et al.), an interesting approach was comparing the sequences obtained by cultivation with those from pyrosequencing (Chapter 2). It is obvious that a larger diversity was to be picked-up by pyrosequencing than by cultivation. In total, sequences could be attributed to 22 phyla for the pyrosequencing, while only five were detected through cultivation. A peculiar observation, however, was when comparing the OTUs from the five phyla in common (*Proteobacteria*, *Actinobacteria*, *Firmicutes*, *Bacteroidetes* and *Deinococcus-Thermus*), that a lack of overlap in the OTUs recovered was noticed. Several explanations were given for the lack in observed overlap between cultivation and pyrosequencing. The most obvious ones are that our sequencing depth – especially to today's norms – is not large enough or that the DNA-extraction was insufficient. Additionally, the cultivation conditions originally used mainly selected for mesophiles and psychrotolerant organisms. Although typical Antarctic bacteria were picked up (e.g., *Flavobacterium micromati*), a part of the cultured organisms may actually represent an exotic and *in situ* non-active fraction of the Antarctic samples, which might only survive temporarily as spores or dormant cells, yet are much faster growers in standard cultivation conditions. Nevertheless, similar observations from the comparison of two techniques, (i.e., a small overlap in taxa obtained) have been made (Donachie et al., 2007; Lagier et al., 2012), and suggests for example methodological aspects such as differences in detection limit (Lagier et al., 2012).

The pipeline used for the pyrosequencing data processing is stringent, and represented the best approach. With hindsight, however, it is regrettable that the RDP reference database was used for the identification of the OTUs. RDP is rather limited in its coverage of bacterial diversity and more and better identifications can be obtained with a more elaborate database

(e.g., GreenGenes). Re-classification in September 2014 indeed showed a significant reduction in unclassified OTUs using the GreenGenes database (data not shown). Although one can argue about the use of classification for short amplicon sequences, a decent identification is important as it may be possible to infer at least some ecological function or information based on the identification of an OTU (but see Koeppel & Wu (2013)).

5.2. HTS data is not used to its full potential.

In a second study, we examined the bacterial diversity and community composition in the Sør Rondane Mountains using the ARISA fingerprinting technique and Illumina MiSeq sequencing. This region was poorly characterized microbiologically, with only two samples subjected to a heterotrophic isolation campaign (Peeters, Ertz, et al., 2011) and two *Cyanobacteria*-focused inventories (Namsaraev et al., 2010; Fernández-Carazo et al., 2012).

The use of fingerprinting techniques allows for a fast and relatively inexpensive screening of samples. Despite certain limitations (e.g., the lack of taxonomic resolution) ARISA has been shown to be robust and reproducible. It allows i.a. for a selection of samples which could then be subjected to HTS. Redundancy analysis showed that the type of bedrock and the presence of macrobiota were important explanatory variables, and that samples generally clustered accordingly. Although ARISA indeed showed to be discriminative (Fig. 3.4b), and can be used in its own right for β-diversity or comparative studies, it would perhaps have been better to have added additional samples to the Illumina sequencing for which sufficient substrate was available, and not limit our scope to those subjected to ARISA. This would have provided more valuable data (sequences), and perhaps would have allowed for a (sub)set of samples for which additional parameters could have been determined.

The amount of unexplained variance was around 80 % and this falls into a broad range of values observed in other studies that work with OTU-level resolution, and is typically for data with a strong presence/absence aspect. It is likely that part of this unexplained variation is due to neutral (random) effects as well as to noise in the data (e.g. sequencing errors and ambiguous signals due to phylogenetic inconsistencies (Koeppel and Wu, 2013)). However, the absence of data for some important nutrients such as nitrogen and phosphorous might be a factor explaining a substantial part of the unexplained variance. In addition, small to microscale differences in topography could, for example, lead to differences in temperature, the amount of solar irradiation, of the frequency of freeze/thaw cycles and local accumulation or depletion of nutrients or elements through run-off. Furthermore, although the bedrock types are generally highly localized and only mix at the margins, mineral composition within a single type is not necessarily homogeneously distributed and fragmentation and a differential leaching of elements will increase the heterogeneity, especially at microscale levels.

HTS studies mainly focus and compare samples on high taxonomic levels (e.g., phylum or class). However, bacteria are intrinsically linked to their (microscale) environment (Vrionis et al., 2013), and there is an increasing demand from the microbiology community to link more organisms to ecology and functions. Except for some groups like *Cyanobacteria,* for most high taxonomic level bacterial groups (i.e., phylum or class), functionality is difficult to deduce and mostly not uniform across lower taxonomic levels. Even at lower taxonomic levels (family, genus and lower), functional capacities of many taxonomic groups remain not or only very partially known. Assessing the importance of biotic and abiotic factors in explaining distribution patterns may provide insights into the structuring role these factors play. This was done for the soil samples (Chapter 3, supplementary data). Nevertheless, because of this intricate relation with the environment, it is important to have sufficient environmental data. The soil study would definitely have benefited from at least additional

information on nitrogen content, which could have supported the hypothesis of e.g. the presence of non-cyanobacterial nitrogen fixers. Additionally, the mineral composition and nutrient release potential of gneiss and granite might also have provided support for our hypothesis that bedrock-type enforces species-sorting. However, logistical restrictions (costs, weight, and preservation during transport) only allowed for a limited number of samples and a small amount of material.

Likewise to the terrestrial sample study, a more complete dataset with environmental variables would have given much more insights into the fine-scale effects of environmental parameters on the bacterial community composition and structure of the microbial mat study. A subset of samples from Victoria Land showed that Ca and SO_4 were the only significant environmental parameters in these samples (data not shown). Nevertheless, it is likely that the environmental parameters available for the complete dataset (pH, salinity and temperature related variables) were the most important ones to explain the patterns on a large spatial scale as similarly shown in studies using for example finger-printing approaches or morphology-based identification of micro-eukaryotes (Verleyen et al., 2003, 2010).

Compared to the terrestrial environment, lakes can be considered to be more confined entities, and it can be hypothesized that organism and nutrient fluxes within a lake are generally larger than exchanges between lakes. More biological replicates of these lake systems would therefore have been very informative to resolve the actual amount of heterogeneity within and between lakes in close proximity. In addition, the number of Sub-Antarctic islands is limited. Because of logistic problems, it was not possible to obtain samples from Crozet Island, located in the South Indian Ocean, which would have been very interesting to compare to Marion Island. Likewise, samples from the South Atlantic Sector would have enabled a more robust view of the Sub-Antarctic.

5.3. Ecological and biological aspects

5.3.1. Environmental parameters affecting bacterial community composition

Our results show an enormous amount of currently unidentifiable organisms, which are potentially new to science. Contrary to temperate or tropical regions where biotic controls influence microbial communities (Cannone et al., 2008; Cary et al., 2010; Stephen D J Archer et al., 2015), abiotic factors are postulated to be the primary influence in polar microbial communities (Hogg et al., 2006; Stephen D J Archer et al., 2015). Both benthic microbial mat and terrestrial communities display a high degree of heterogeneity at local (regional) and continental scales, suggesting the importance of both large-scale and microscale environmental factors on the community structures. It has been suggested and generally accepted that, at least for soil bacteria, pH is the most forcing variable in bacterial biogeography and diversity (e.g., Fierer and Jackson 2006; Griffiths *et al.* 2011; Feng *et al.* 2014). In our studies, pH indeed appeared to be an important explanatory variable, although for the microbial mat samples, temperature and conductivity were more important. A similar observation was made for meltwater ponds in the Ross Sea Region, where – although not unexpectedly for this type of habitats – conductivity was the main environmental driver, and pH showed only a poor correlation (Stephen D J Archer et al., 2015). For soils, carbon availability was more important. In a global, non-Antarctic drylands study, increasing aridity had a negative impact on bacterial diversity, mainly as the result of a reduced carbon content (Maestre et al., 2015). Because of the oligotrophic and extreme climatologic conditions in the Antarctic, bioavailable water indeed is the limiting factor for life and microbial activity, and together with organic matter also appears to be an important driver of bacterial diversity and community structure. A higher carbon content correlated i.a. with a higher proportion and richness of *Alphaproteobacteria,* which was for example also demonstrated to be the case in

rhizospheres (e.g. Gottel et al., 2011), while a higher water content increased archaeal richness in MDV soils (I. Richter et al., 2014). Macrobiota are highly localized sources of organic matter, while additionally, plant cover is likely to mitigate the climatic conditions (Yergeau, Newsham, et al., 2007). The lack of buffering in Antarctic ecosystems probably accounts for a part of the observed spatial heterogeneity, since it increases the possibility of highly localized (microscale) environmental conditions and effects.

Conductivity appeared to be the main environmental driver for the continental microbial mat samples, but also significantly impacted soil communities. For example, *Chloracidobacteria* showed a high relative abundance in both terrestrial and aquatic samples with a low conductivity. Besides EC, pH and nutrient content (i.e., carbon and nitrogen), the elemental composition (e.g., Fe, Pb, Mb, V, …) of soil or lacustrine samples can be vary locally, and will significantly impact bacterial community structure (Stephen D J Archer et al., 2015; Kim et al., 2015) as they provide necessary elements for several bacterial functions, such as photosynthesis or nitrogen fixation.

5.3.2. Biogeography and macroecological patterns

It was long thought that because of their sheer abundance and minute size, microorganisms did not show any dispersal limitation and hence displayed no biogeographical patterns since geographical distance was irrelevant (i.e., isolation by geographic distance) (Fenchel and Finlay, 2004). Accumulating evidence, however, shows that this is not the case, with geographic barriers effectively isolating microbial populations (Cho and Tiedje, 2000; Whitaker et al., 2003). Advances in genetic technologies have enabled the existence of microbial biogeography. There is now evidence that microorganisms display patterns similar to macroorganisms, such as distance decay, abundance-range relationship and Rapoport's rule

(Amend et al., 2010; Hanson et al., 2012). Although the microbial mat samples showed a high dissimilarity, only a weak – yet significant – distance decay (not shown) was observed, suggesting that environmental forcing is stronger that distance effects. In addition, a study on marine surface bacteria showed similar bacterial communities at both poles (Amend et al., 2013).

From our study on microbial mats (Chapter 4), it appears that temperature is an important variable distinguishing Sub-Antarctic samples from the Antarctic mainland (Fig. 4.5), while electric conductivity was important within the continent and regions.

5.3.2.1. Isolation of Antarctica and dispersal capacities of bacteria

It is generally assumed that the influx of organisms is largely prohibited by circum-Antarctic wind or oceanic currents (Chong et al., 2015). Studies further suggest that dispersal occurs mainly over short distances (e.g., Wood et al., 2008). However, relatively few studies have focused on Antarctic aerobiology (e.g., Pearce *et al.* 2009, 2010), and recently, a consortium was established to raise the awareness of this lacuna in order to enable comparative studies (Pearce et al., 2016). Although the sheer number of OTUs prohibits the manual comparison (BLAST) of every single OTU, a BLAST search of the most numerous or most widely distributed (shared between samples) OTUs was performed. Often, such queries resulted in sequences having 98 – 100 % identity with sequences from hot deserts. This suggests that the ability to cope with drought is a common trait of these organisms. However, because of the supposedly intrinsically low evolutionary rate and short amplicon length giving relatively low resolution, based on our data, it cannot be concluded that an exchange or influx of organisms readily occurs.

5.3.2.2. Latitudinal gradient in OTU-richness

Latitudinal diversity gradients have been observed for macroorganisms (plants and animals) for over two centuries, with generally an increasing species richness from the poles to the (sub)tropics (Fine, 2015). Only recently have genetic approaches enabled for similar patterns to be recognized unambiguously for microorganisms, for example in bacteria s.l. (Fuhrman et al., 2008) and at genus level (Andam et al., 2016).

Two not entirely mutually exclusive ecological mechanisms are considered to be at play, namely productivity and environmental temperature. A latitudinal diversity gradient for marine planktonic bacteria has been established with different methods, and correlated mainly to (average annual) sea-surface temperature (Fuhrman et al., 2008). In oligotrophic marine regions, a positive energy-diversity relationship was found, which levelled off in mesotrophic conditions (Bienhold et al., 2012). Although our samples covered a large range in environmental parameters – some of which might affect richness, such as conductivity –, the gradient was still valid for samples with a comparable EC.

The observation that Antarctic climatic and environmental conditions are extreme can largely be considered an anthropogenic view. Organisms active in or endemic to Antarctica, or other 'extreme' environments, have been adapted to these conditions. Indeed, psychrophilic organisms are performing badly under mesophilic conditions. Nevertheless, low temperatures affect water activity (a_w) and hamper even organisms adapted to these conditions, with a reported lower limit of -20 °C for bacterial activity (Margesin and Miteva, 2011).

> As a side note, reflecting on this temperature, one could wonder what the effects are on the microbial community structure or the environmental conditions for (Antarctic or polar/low temperature environment) samples which are (temporarily) stored at -20 °C, a standard preservation or transportation temperature.

The low temperature – either as a direct effect or through the limitation of biologically available water –, combined with a low energy availability and/or a high UV irradiance, likely explains at least part of the observed relatively low richness on the continent, and prevent a fast adaptation of exotic organisms. On the other hand, it was suggested that a high degree of HGT could occur between organisms in hypolithic environments (Cowan et al., 2011). Alien species might hence be able to establish themselves in such environments.

5.4. Technical aspects

Processing HTS data, or any (ecological) data, requires numerous decisions to be taken, such as algorithms or programmes to be used or the definition of cut-offs (e.g., for OTU-clustering, or the minimum abundance and/or distribution an OTU should have (i.e., removal of singletons, doubletons, ...)). I applied a relatively stringent quality control and used standardized procedures throughout my thesis. An overview and short discussion on the advantages this stringent control as well as on some technical aspects is given below.

5.4.1. Mock-communities

For the studies described in Chapters 3 and 4, we switched from Roche's pyrosequencing to the Illumina MiSeq platform, providing us with more and longer sequences. To aid in the decisions to be made during processing and in order to monitor the quality of the data within and between runs, mock-communities were incorporated. Twenty-one OTUs from the pyrosequencing data that were available through the cultivation efforts of Dr. Karolien Peeters (Peeters, Ertz, et al., 2011; Peeters, Hodgson, et al., 2011; Peeters, Verleyen, et al., 2011) were selected, comprising the five cultivable phyla (*Actinobacteria*, *Deinococcus-Thermus*, *Proteobacteria*, *Bacteroidetes* and *Firmicutes*) (see Addendum II). Extracted DNAs were pooled equimolarly and consequently treated as a sample during the library preparation. This mock-community was included several times in each run. A provisional attempt to process

Illumina data was done according to MOTHUR's Illumina pipeline (version of January 2015). However, this left us with over 200 OTUs in the tested mock community. Using USEARCH (Edgar, 2010) for sequence filtering and UPARSE (Edgar, 2013) for OTU-clustering, this number was reduced to 20 and 21 OTUs (19 shared) for the Chapter 3 data and to 25-32 OTUs for the Chapter 4 data. Overall, 20 OTUs were shared between the five mock communities from two sequencing runs (three in the first run and two in the second run). In all cases, the two most abundant OTUs each made up about 15 to 20 % of the reads, while the others gradually but consistently declined between mock communities within a study from ~10 % to ~1.5 to 0.1 % of the shared OTUs (in both studies a *Deinococcus* sp.). Several additional OTUs could be detected at very low numbers (mostly singletons) within a mock-community, of which some were shared between mock-communities of different runs and apparently were related to the more abundant OTUs (thus being either errors or variants), although few were clearly contaminants (especially cyanobacterial singletons).

While pipetting errors may have contributed to these observations, they likely also reflect the bias associated with PCR-based approaches and call for prudence when considering abundances. This clearly indicates the necessity to add controls, such as mock-communities, negative controls and replicates, to these kinds of studies in order to adjust parameters in the sequence-processing steps, although, of course, it remains unclear to what extent this can be extrapolated to environmental samples. Mock-communities and replicated samples each clustered together, indicating consistency and reproducibility. In addition, the dominance of for example *Cyanobacteria* and *Alphaproteobacteria* in microbial mats (Chapter 4) and the correlation of *Cyanobacteria* and *Proteobacteria* with high moisture content and organic matter, and *Acidobacteria* and *Actinobacteria* with low moisture content in soils (Chapter 3), does seem to be legitimate (Varin et al., 2012; de los Ríos et al., 2015).

5.4.2. Removal of extracellular DNA

In order to allow comparison with older datasets from both laboratories and to be able to compare results between several ongoing studies, we have used the same extraction protocols throughout our different set-ups. This included an initial step of removal of extracellular DNA (eDNA) as described by Corinaldesi and colleagues (2005), which was primarily done in order to obtain a more correct view on the active eukaryote community members. This is particularly interesting when trying to correlate environmental parameters with the sequencing data. However, this might result in a distorted perception of the actual active organisms, since, for example, DNA from dormant, recently deceased, or (endo)spores can be extracted (Blazewicz et al., 2013; Cangelosi and Meschke, 2014). Nevertheless, a sample represents only a snapshot in time, and might for example be just one stage in a (cyclic) succession, which is not unlikely in the Antarctic (Brinkmann et al., 2007). eDNA may hence originate from dead community members that were at some earlier time active and highly abundant, and are potentially able to thrive in the same physical and chemical settings, but for instance under different microclimatic conditions. Removing all eDNA thus potentially thwarts insights into probably important players that no longer have (many) active cells under current environmental conditions.

5.4.3. The 16S rRNA gene variable region used

We have studied the V1-V3 variable regions of the 16S rRNA gene as this appears to be the most informative region on different taxonomic levels (Jeraldo et al., 2011; Yarza et al., 2014). V1-V3 is also most congruent with PCR-independent metagenomics and very good at detecting *Chloroflexi* although it performs worse for certain gammabacterial groups. The V4 region, as recommended by the Earth Microbiome project (Knight et al., 2012) performs well in silico, but greatly underestimates the *Actinobacteria* and *Chloroflexi* (Albertsen et al.,

2015). These are two groups which are important components of both soils and aquatic systems, as was also shown from our studies. We therefore consider the V1-V3 region a good choice, although several other studies have used other variable regions. When comparing studies this aspect needs to be taken into account.

In addition, it also seems that preferential amplification in high diversity (i.e., environmental) samples is less likely with our primer set and our moderate (55 °C) annealing temperature, as was shown by Sipos et al. (2007) and Albertsen et al. (2015). Indeed, no single OTU was consistently numerically dominant in every sample, which could be expected to be the case if amplification would be severely biased, as more or less observed in the mock-communities.

5.4.4. OTU cut-off

An integral aspect of amplicon-based sequencing is the use of OTUs. Although it was demonstrated in Chapter 2 that the default use of a single "one cut-off fits all" 97 % (or similar) value for any 16S rRNA amplicon length or variable region or taxon is incorrect, this practice currently allows for comparison with other studies. It should also be noted that there are a plethora of binning algorithms, and that there is no consensus for which one to use. This too results in different perceived richness values (Powell and Sikes, 2014; He et al., 2015).

5.5. Perspectives

It is striking how relatively few studies have been performed on the Antarctic and Sub-Antarctic islands. For example, no studies on the microbial diversity of Macquarie Island are available (except for mouse pathogens (Moro et al., 2003)). While the number of Antarctic microbiological studies is rising, few have actually been based on the 16S rRNA gene (Chong et al., 2015), and there have been no large-scale, region-wide systematic surveys of the bacterial diversity, community structure and distribution. This is however urgently needed, given that (i) many regions and habitats (e.g. epiglacial lakes and cryconites) are still unexplored or under-sampled, (ii) the incidence of endemism appears to be high, (iii) some sectors of Antarctica are amongst the most rapidly warming regions on Earth (IPCC 2013), and (iv) habitats in these regions appear to respond quickly to climate changes (e.g., Quayle et al., 2002). Of course, the vast size and inaccessibility of the continent make this an extremely difficult task, not to mention the costs associated. However, to reduce the variation inherent with HTS data, and to allow for comparison (Jones et al., 2015), consortia have been instigated to promote a uniform approach toward sampling and the processing hereof (e.g., the Earth Microbiome Project), and this can be applied to Antarctic environments within for example working groups of the Scientific Committee for Antarctic Research (SCAR). However, to increase the comparability of data between regions, not only should the sampling itself be performed in a uniform way, also a well-defined sampling design should be adhered to (see below).

Besides the extent of the total bacterial richness in Antarctica, many other questions remain unanswered such as, (i) What is the functional diversity, and more important, who's active and how do the microorganisms function *in situ*? (ii) What is the overlap of bacterial taxa between different environments such as soils and lakes and are there source-sink dynamics

between these habitats? (iii) How do they interact (within and between the domains of life)? (iv) How does community structure and functionality change within and between seasons and between years? (v) How and to what extent do microorganisms disperse within the continent? (vi) What is the rate of exchange with other regions, and what proportion of microorganisms is endemic? (vii) How do the Antarctic microbial communities compare to those in similar Arctic and alpine environments, and what is their phylogenetic relationship? (viii) How and where did polar microorganisms survive during glacial maxima, does this provide information about their current distribution and how is this related to the spatial distribution of multicellular organisms?

5.5.1. Suggestions for future research

Several of the questions mentioned above can already be (at least partially) examined. Using the sequences already available from our soil and microbial mat samples, we can look at the fine-scale phylogenetic diversity. This can for example provide insights in whether aquatic and soil bacteria are distinct phylotypes, and, whether sequences that are binned into OTUs are more closely related within a region or habitat than with those from in other regions or habitats (Koeppel and Wu, 2013). Adding sequences from public databases would allow for a more extensive comparison and can reveal to what extent interhemispheric dispersal has occurred between the Arctic and Antarctica. Also, potential colonisation events from the other Southern Hemisphere continents into the Antarctic can be potentially detected using a more global-scale comparison, provided that the dataset is sufficiently large for similar habitats in these regions.

In addition, a phylogenetic analysis would also provide more resolution for the sequences currently unclassified at phylum level.

Besides amplicon-based biodiversity studies, HTS approaches are now sufficiently established to be used in well-designed, replicable experiments too (Prosser, 2015). In a first step, several interesting soil and lacustrine samples could be selected based on our amplicon sequencing approach and subjected to a shotgun metagenomics survey. This will allow uncovering the functional potential of microbial communities of a wide environmental range from different regions. To date, only a handful of such studies are available for the Antarctic (e.g., López-Bueno et al., 2009; Ng et al., 2010; Lauro et al., 2011; Yau et al., 2013) and provided a first insight into the functioning of microorganisms in the extreme conditions. This trait-based approach also has the potential to refine the perceived biogeographical patterns (Green et al. 2008), since the primarily taxonomical approach used might not provide enough resolution.

Furthermore, most studies only look at one particular component of a microbial community: *Cyanobacteria*, heterotrophs, fungi, unicellular eukaryotes, or even subgroups (e.g., a single genus). However, microorganisms in a community interact, and ecosystems can only be truly understood when the entire community is considered. *Cyanobacteria* are thought to be the primary carbon fixers, yet, in certain samples (Chapter 3) they appeared to be absent. Although other bacteria are able to fix carbon, they do so at much lower rates. However, considering only bacteria ignores the possible contribution of microalgae, such as diatoms. Hence, a shotgun metagenomics or –transcriptomics approach will provide insights into actual community functioning and identify key those organisms which were not targeted using universal primers for bacteria alone. Moreover, in order to fully understand the effect of Climate Change (Cavicchioli, 2015) on these ecosystems, biodiversity assessments should be combined with functional genomics approaches and measurements of their ecosystem functions (Adams et al., 2006; Temperton and Giovannoni, 2012).

In combination with these culture independent approaches, growing the organisms in the lab will remain necessary in order to truly understand the physiology of these bacteria and their roles in ecology and natural product production (Stewart, 2012). Large-scale cultivation under a broad range of growth conditions (culturomics) and for example the use of gellan gum instead of agarose (Tanaka et al., 2014) will increase the number of organisms that can be cultured. In addition, extracting information from metagenomics data might provide information on how these organisms could be cultured (Giovannoni and Stingl, 2007). Two taxa encountered in this thesis are in this respect of particular interest to focus cultivation attempts on, namely *Acidobacteria* and the candidate division *FBP*. *Acidobacteria* are thought to be ecologically important in (terrestrial) environmental samples and show a high diversity (Quaiser et al., 2003; Jones et al., 2009). However, cultivation and isolation proves difficult with few successes (Ward et al., 2009). The high relative abundance of *Chloracidobacteria* in certain samples could warrant such efforts. The negative correlation of Antarctic *Chloracidobacteria* with conductivity and total organic carbon content (Chapter 3) already provides a clue to compose growth media with nutrient-poor conditions (low or no additional carbon) and low ion or salt concentrations. Using a long incubation time at low temperatures in combination with different pH ranges and light regimes could pay off.

In addition, the candidate phylum *FBP*, closely associated to the phylum *Armatimonadetes* (Lee et al., 2013) is of particular interest, because this phylum was first discovered from Antarctic samples. Moreover, environmental sequences identified as belonging to this phylum are mainly associated with polar and alpine samples, and were indeed (nearly) absent in our Sub-Antarctic samples (Chapter 4). Cultivation itself might provide material to investigate the genomic peculiarities of a most likely truly cold-associated phylum. The availability of environmental sequences from databases, however, can already be used to conduct a

phylogenetic survey of this phylum to investigate its biogeography, which could provide evidence for dispersal abilities, as also mentioned above.

5.5.1.1. A suggestion for controlled experiments

It can be expected that much novel diversity is still to be discovered. However, it remains unclear to what extent bacteria and other microorganisms are widely distributed. Per definition, studies are limited in spatial resolution, both at the microscale and at larger spatial scales, and in the number of samples. Studies either adhere to a cherry-picking (i.e., random) sampling strategy or have a limited spatially explicit design. In addition, random sampling could potentially inflate β-diversity (Zhou et al., 2013) . The apparent observed large spatial heterogeneity could therefore be smaller when a spatially explicit set-up is used, i.e. much of the diversity could be retrieved when fewer locations would be sampled but more intensely. For example, samples taken 50 cm apart in a square grid of 25 m² could be repeated several times within a region. Replicates should be taken at every sampling point to try to incorporate microscale heterogeneity. The use of such *in situ* experimental set-ups will allow for decent monitoring of community turnover along gradients of e.g. moisture and TOC or macrobiota load, and time. In combination with for example temperature manipulation through open top chambers in several evenly spaced sub-plots, this would suit particularly well for a metatranscriptomics approach, which is more informative than mere metagenomics (Prosser, 2015). Such experimental set-ups could then be employed in different regions of the continent, allowing genuine comparisons, e.g. about the potential differences in the effects of climate change in different regions, on different soil types and under different environmental conditions (e.g. dry vs. moist soils or mineral soils vs. macrobiota containing soils).

The older microbial mat samples were in first instance collected to investigate eukaryotic microorganism diversity, and hence the environmental parameters available are attuned to

such data. As bacteria are apparently much more sensitive to the microenvironment, local elemental composition needs to be determined in order to increase our understanding of the correlations with and reactions of bacteria on the environment. For future studies, a standardized list of minimal environmental parameters that need to be determined is hence necessary, for example, based on the MIMARKS set (Yilmaz et al., 2011; Knight et al., 2012).

5.5.2. Technological advances

Molecular techniques are evolving fast. Currently, the largest concerns are the retention of errors and the (still) relatively short length of the amplicon reads. Although the arrival of new approaches, such as nanopore sequencing (Maitra et al., 2012), is expected to deliver much longer reads (potentially 10s kb) at a much lower error rate (Laver et al., 2015), as with any new technology, assessments of its performance are needed and new approaches need to be developed to deal with the earlier problems (Check Hayden, 2014; Mikheyev and Tin, 2014; Jain et al., 2015; Laver et al., 2015). On the other hand, improvements to existing methods have been and are being developed to reduce the error rate, such as low-error amplicon sequencing (Faith et al., 2013). Increased read lengths and decreased error rates will bring more robust estimates for α and β-diversity, and will allow for more correct taxonomic and phylogenetic inferences.

Additionally, bioinformatics solutions and improvements are being developed even faster, and allow for a further reduction of errors and improved diversity estimates. Recently, a new approach, oligotyping, was developed. This approach enables the detection of more fine scale ecological patterns for microbial communities through the comparison of closely related sequences (Eren et al., 2013, 2015). Such improvements will be particularly useful for the further analysis of biogeographical patterns, especially when comparing global cold regions, since a lower evolutionary rate urges for more sensitive approaches.

However, since such surveys are still amplicon-based, they lack the resolution of the comparison of complete genomes, which ultimately is the only way to really distinguish organisms. A trait-based approach could for example reveal finer biogeographical or temporal patterns than observed purely by taxonomy (Green et al. 2008). Currently, metagenomics, or genomics for those organisms that can be isolated, is the most informative, although in the longer term, deep single-cell sequencing might be become the default approach.

To reduce the number of unclassified reads, or more correctly, to have an idea of which sequences have been detected in Antarctica before, a reference training set used by a classifier algorithm could be built containing the unique (non-redundant) sequences detected in Antarctica. In a second phase, this could be extended with Arctic and alpine sequences. This way, it would also be possible to see in which regions a particular sequence has been discovered before.

In summary, the proposed research both includes (i) a larger sampling coverage, in region, habitat and at the microscale and (ii), the use of novel techniques enabling a more accurate and complete assessment of the richness, community structure of organisms from all domains of life as well as their functional genetic make-up.

5.6. General conclusions

In the last decade, the use of HTS has enabled the in-depth genetic characterization of bacterial communities, providing an idea of bacterial diversity and enabling to detect biogeographical and macroecological patterns. In this thesis, we showed that, while local diversity is relatively low on mainland Antarctica compared to lower latitude Sub-Antarctic islands, the overall diversity is still large with many unknown taxa, which may potentially be endemic. A latitudinal gradient in OTU richness was observed, which is likely due to a combination of isolation (dispersal limitation) and temperature constraints.

Temperature also appears to be the major variable explaining large scale biogeographic patterns in littoral microbial mats, i.e., between Sub-Antarctica and the Antarctic mainland. Within the continent between regions, geographic isolation and – correlated – possibly to a certain extent dispersal limitation. Within regions, conductivity is arguably the main environmental parameter differentiating communities, while other physicochemical parameters such as nutrients (C, N, P), pH, oxygen concentrations and elements, will further have a fine scale effect both within and between lakes. Similar conclusions are likely apparent for terrestrial communities, although because of the lack of buffering, effects are much more localized and likely much more pronounced (e.g., nutrient or salt concentrations).

This thesis hence supports the hypothesis that for Bacteria in the Antarctic Region, too, the environment exerts large selective pressures on community structure and composition, complemented by biotic factors. Although we only were able sample a fraction of the continent, it is expected that similar patterns hold across the entire continent. However, additional sampling and in depth (metagenomics) sequencing linked to extensive environmental data, combined with phylogenetic analysis is needed to resolve important

questions such as within and inter-continental dispersal, functioning and correlation of observed patterns to environmental data.

Summary

Geological events and historical climate changes have eliminated or reduced most life in Antarctica to mainly microbial organisms in relatively simple communities. Due to its exceptional location, millennia long isolations and extreme climatic conditions, the continent offers a spectacular and unique background for fundamental scientific research and the testing of hypotheses. Notwithstanding the fact that Antarctica is still considered by many to be one of the last pristine environments on Earth, it is not only threatened by climate change, which particularly has severe effects on parts of West and Maritime Antarctica, but also by an ever increasing number of tourists and even scientists themselves.

Studies on Antarctic biota are relatively scarce, and despite the fact that bacteria are fundamental to the Antarctic ecosystems, only a minority of the studies focus on these organisms. This results in a lacuna in the knowledge about the diversity, distribution and functioning of and the relationships between these organisms under the extreme Antarctic conditions. The recent advent of High-throughput sequencing (HTS) applications enables to sequence millions of DNA-fragments in a very short time, allowing us to visualise bacterial communities at a very high resolution, without the necessity for prior isolation of the organisms. In this PhD-study, we have applied some of these new technologies in order to investigate the bacterial diversity of different habitats throughout the Antarctic.

In a first study (Chapter 2), we have compared the results obtained by pyrosequencing and compared these with the results of a previous isolation campaign. As expected, a much larger diversity of bacteria were found with pyrosequencing. While five bacterial phyla were recovered by cultivation, this was the case for 22 phyla with the NGS-approach, and a large amount of unknown diversity was evident. At the same time, it became clear that also the part of the 16S rRNA gene that was sequenced had an impact on the perceived diversity, with the

V1-V2 fragments resulting in ~50% more OTUs than the V3-V2 fragments and only a limited amount of overlap in the genera recovered was noticed. In contrast, more chimeric sequences were identified in the V3-V2 amplicons. Notwithstanding the fact that pyrosequencing yielded a higher diversity, there was very little overlap with the cultivation approach, with only about 4 % of the OTUs recovered by cultivation found with pyrosequencing. In contrast, we also noticed that some singleton pyrosequencing OTUs where easily grown on growth media, and hence were not errors in the pyrosequencing data. This study thus showed that several factors could have a large impact on the perceived diversity, and that complementary techniques are necessary to discover the total bacterial diversity.

In a second study (Chapter 3), we have examined the effects of both different bedrock types (granite and gneiss) and the presence of macrobiota (mosses, lichens and algae) on the composition of bacterial communities in high-altitude inland soils of different regions if the western Sør Rondane Mountains (Queen Maud Land, East Antarctica), near the Belgian Princess Elisabeth research station. We have used the at present most used HTS-platform, Illumina's MiSeq, which allows sequencing longer gene fragments and yields more sequences compared to pyrosequencing. We combined this with the ARISA genetic fingerprinting technique. We demonstrated that organic carbon was the most significant parameter in structuring bacterial communities, followed by pH, electric conductivity, bedrock type and moisture content, while spatial distance was of less importance. Diversity showed a positive correlation with moisture content. *Acidobacteria* and *Actinobacteria* dominated dry gneiss derived mineral soils, while *Proteobacteria*, *Cyanobacteria*, *Armatimonadetes* and candidate phylum *FBP* were dominant in samples with a high organic carbon content. A large part of the unexplained variation is probably caused by the absence of data about important nutrients in our dataset (nitrogen and phosphorous), together with microclimatic and topographic differences between sample locations, and noise and stochasticity.

In a last study we again used the Illumina MiSeq platform to perform a pan-continental charting of benthic and littoral microbial mats. In total, 138 samples from lakes in eight Antarctic regions and two Sub-Antarctic islands were analysed. We found a significant trend of increasing biodiversity with decreasing latitude from 85° to 54° S, although this than again decreased until 45° S. The mean annual temperature appeared to have a highly significant effect on community structure between Sub-Antarctica and Antarctica, while, besides the geographical distance, electric conductivity, and to a lesser extent pH, was important in explaining differences between samples on the Antarctic continent. In this study, too, a very high unknown diversity was observed. Particularly *Cyanobacteria* and *Alphaproteobacteria* dominated freshwater microbial mats, while *Bacteroidetes* and the alphaproteobacterial *Rhodobacteraceae* family dominated saline lakes. The Sub-Antarctic Marion Island was highly deviant with very low species richness, dominated by *Janthinobacterium* (*Betaproteobacteria*).

In conclusion this thesis supports the hypothesis that for Bacteria in the Antarctic Region, too, distinct biogeographic patterns exist and that the environment exerts large selective pressures on community structure and composition, complemented by biotic factors. There is a high amount of heterogeneity at both local and continental scale due to both spatial distance and local differences in environmental variables such as electric conductivity, pH, moisture content, organic carbon and microclimate. Although we only were able sample a fraction of the continent, it is expected that similar patterns hold across the entire continent. However, additional sampling and in depth (metagenomics) sequencing linked to extensive environmental data, combined with phylogenetic analysis is needed to resolve important questions such as within and inter-continental dispersal, functioning and correlation of observed patterns to environmental data.

Samenvatting

Geologische gebeurtenissen en historische klimaatsveranderingen hebben het meeste leven op Antarctica geëlimineerd of gereduceerd tot in hoofdzaak microbiële organismen in relatief eenvoudige gemeenschappen. Door de uitzonderlijke ligging, millennia-lange isolatie en extreme klimaatcondities, biedt het continent een spectaculaire en unieke achtergrond voor fundamenteel wetenschappelijk onderzoek en het testen van hypothesen. Hoewel Antarctica voor velen nog als een van de laatste ongerepte plekken op Aarde geldt, staat het niet enkel onder druk van de klimaatsverandering, die in het bijzonder delen van West en Maritiem Antarctica hard treft, maar ook in toenemende mate van toerisme en de onderzoekers zelf.

Studies over de Antarctische biota zijn relatief zeldzaam, en hoewel bacteriën aan de basis liggen van de Antarctische ecosystemen, handelt slechts een kleine minderheid van de studies over deze organismen. Dit resulteert in een leemte in de kennis over de diversiteit, de verspreiding en het functioneren van en de onderlinge relaties tussen deze organismen in de extreme Antarctische condities. De recente komst van Next Generation Sequencing (NGS) toepassingen laat toe om miljoenen DNA-fragmenten te sequeneren op een kort tijdsbestek, waardoor het mogelijk geworden is om de bacteriële gemeenschappen op een zeer hoge resolutie te aanschouwen, zonder de noodzaak om deze organismen vooraf te isoleren. In deze doctoraatsstudie hebben we enkele van deze nieuwste technologieën aangewend om de bacteriële diversiteit in verscheidene habitats verspreid over het Antarctische continent na te gaan.

In een eerste studie (Hoofdstuk 2), hebben we de resultaten van pyrosequencing vergeleken met de resultaten van een voorgaande heterotrofe isolatiecampagne. Zoals verwacht, werd een grotere diversiteit aan bacteriën gevonden met pyrosequenering. Terwijl vijf bacteriële fyla waren gedetecteerd door cultivatie, waren er dit 22 met de NGS-techniek, en werd een grote

hoeveelheid ongekende diversiteit aangetroffen. Tegelijkertijd werd het duidelijk dat ook het deel van het 16S rRNA gen dat gesequeneerd werd een impact had op de waargenomen diversiteit, waarbij de V1-V2 fragmenten ~50% meer OTU's opleverden dan de V3-V2 en er een beperkte overlap was in de opgepikte genera voor beide fragmenten. Daarentegen werden in de V3-V2 amplicons een relatief groter aantal chimeren aangetroffen. Niettegenstaande het pyrosequeneren meer diversiteit kon oppikken, bleek er toch weinig overlap te zijn met wat via cultivatie bekomen was, en werd slechts een 4-tal procent van de cultivatie OTU's teruggevonden via pyrosequenering. Anderzijds konden we ook vaststellen dat OTU's die maar eenmaal opgepikt werden (singletons), soms heel gemakkelijk te groeien waren op groeimedia, en dus geen fouten waren in de sequentie data. Deze studie toonde dus aan dat verschillende factoren een grote invloed kunnen hebben op de waargenomen diversiteit, en dat het gebruik van complementaire technieken noodzakelijk is om de volledige bacteriële diversiteit in kaart te brengen.

In een tweede studie (Hoofdstuk 3), hebben we het effect nagegaan van zowel verschillende ondergronden (graniet en gneiss) als de aanwezigheid van macrobiota (mossen, korstmossen en algen) op de samenstelling van bacteriële gemeenschappen in hoger gelegen bodems in verschillende regio's van de westelijke Sør Rondanebergen (Koningin Maudland, Oost Antarctica), nabij het Belgische Prinses Elisabeth-onderzoeksstation in het Antarctische binnenland. Hierbij hebben we gebruik gemaakt van het huidige meest gangbare NGS–platform, Illumina MiSeq, dat het mogelijk maakt om zowel langere stukken van genen te sequeneren, alsook om meer data te genereren dan mogelijk was met pyrosequenering, in combinatie met de genetische fingerprinting-techniek ARISA. Hierbij werd aangetoond dat organische koolstof de meest bepalende factor was voor de bacteriële gemeenschapsstructuur, gevolgd door pH, elektrische geleidbaarheid, ondergrond en vochtinhoud, terwijl ruimtelijke afstand minder belangrijk was. Diversiteit toonde een positieve correlatie met vochtinhoud.

Acidobacteria en *Actinobacteria* domineerden de droge gneissgebaseerde minerale bodemstalen, terwijl *Proteobacteria*, *Cyanobacteria*, *Armatimonadetes* en kandidaat-fylum *FBP* dominant waren in stalen met een hoge organische koolstofinhoud. Een groot deel van de onverklaarde variatie is waarschijnlijk te wijten aan verschillen in microklimaat tussen de verschillende stalen.

In een laatste studie werd opnieuw van het Illumina MiSeq-platform gebruik gemaakt om een pancontinentale kartering te doen van benthische en littorale microbiële matten. In totaal werden 138 stalen geanalyseerd, afkomstig van meren in acht Antarctische regio's en twee Sub-Antarctische eilanden. Er werd een significante trend gevonden waarbij de biodiversiteit toenam met afnemende breedtegraad van 85 tot 54° Zuid, om daarna terug af te nemen tot 45° Zuid. De gemiddelde jaartemperatuur bleek een sterk significant effect te hebben op de gemeenschapsstructuur tussen Sub-Antarctica en Antarctica, terwijl, naast de onderlinge geografische afstand, elektrische geleidbaarheid belangrijk was in het verklaren van de verschillen tussen stalen op het Antarctische continent. Ook in deze studie werd een zeer hoge ongekende diversiteit waargenomen. Voornamelijk *Cyanobacteria* en *Alphaproteobacteria* domineerden de microbiële matten van zoetwatermeren, terwijl *Bacteroidetes* en de alfaproteobacteriële familie *Rhodobacteraceae* de zoutwatermeren domineerden. Het Sub-Antarctische Marion Island was erg afwijkend met een zeer lage soortenrijkdom gedomineerd door *Janthinobacterium* (*Betaproteobacteria*).

Samenvattend kan gesteld worden dat op basis van deze thesis duidelijk is dat er ook voor bacteriën in Antarctica duidelijke biogeografische patronen bestaan, en dat de omgeving een grote selectiedruk uitoefent op de structuur en samenstelling van bacteriële gemeenschappen, met daarbovenop biotische factoren. Er blijkt dat er een hoge mate van heterogeniteit op zowel lokale als continentale schaal, te wijten aan zowel ruimtelijke afstand en lokale verschillen in omgevingsvariabelen zoals elektrische geleidbaarheid, pH, vochtinhoud,

organische koolstof en microklimaat. Alhoewel we slechts een fractie van het continent hebben kunnen bemonsteren, kunnen we verwachten dat gelijkaardige patronen voor het gehele continent geldig zijn. Bijkomende staalnames en metagenoomstudies die gelinkt kunnen worden aan een uitgebreide dataset, gecombineerd met fylogenetische analyses, zijn nodig om belangrijke vragen op te lossen, zoals intra –en interconinentale verspreiding, het *in situ* functioneren en het correleren van de patronen met omgevingsdata.

Addendum I: Technical recommendations for future research

I.1. Wet lab

- In order to reduce biases (preferential amplification, …) keep the number of PCR-cycles as low as possible. Dependent on the initial concentrations, this should preferably be between 20 and 30 cycles.
- Avoid long stretches of single adapters-barcode-primer oligos, as they could promote autoligation. A two-step PCR approach might be better (Berry et al., 2012), which is per definition the case when using e.g. the Illumina Nextera kits.
- Use high fidelity and hot-start DNA polymerases. This will reduce erroneous nucleotide insertions and PCR-mediated recombination (chimeras).
- Increase the elongation time (at least 90 s to up to 5 min) in order to let the polymerases complete the elongation (Engelbrektson et al., 2010; Consortium et al., 2012). This will increase the yield per cycle and will reduce chimera formation by reducing the amount of uncomplete template which could then serve as a primer.
- Primers should be HPLC-purified; This is not recommended for degenerated primers, as this might disrupt the concentrations of the different primers. It should be noted, however, that the use of degenerated primers themselves could potentially bias PCR (Vos et al., 2012).
- Be aware for cross-contamination through microdroplets/spray when doing library preparation. Include negative controls.

I.2. Sequence processing

- Include mock communities. These may be used to calibrate processing parameters, check run quality, check for contamination, and consistency between runs. These may

be very general and of a broad taxonomic composition, or may be tailored to what is expected in the samples. For a broad scope, it is interesting to include several phyla and different lineages within a phylum, and both Gram-positive and Gram-negative organisms. In either case, to check for resolution of the sequencing and consequent sequence processing, (two or three) closely related strains (checked *a priori* to effectively result in different OTUs) from two or more genera might be included. Most mock communities consist of about 20 specimens, but more may be more robust to PCR biases, and more likely represent natural communities.

- Similarly, always include a specific (well-characterized) sample with each run, which can be used to control for consistency (a so-called generous donor sample).
- Perform *de novo* chimera checking within individual samples, and not based on the complete sequence pool, since chimeras are formed within a single PCR reaction. Avoid closed reference-based chimera checking (e.g., using the Silva Gold Reference database), since this could possibly lead to false positives when a lot of novel diversity is present for which no representatives are present in the reference. Currently, Uchime (Edgar et al., 2011) appears to be performing best, with similar results obtained by Perseus (Quince et al., 2011), although the latter is much slower.
- Include a mixture of the PCR negative controls. This is especially useful when sequencing bacteria. In environmental samples, try to identify human symbionts (e.g., *Corynebacterium*, ...), when these are not relevant for the research question. However, examine distribution and abundance prior to throwing out OTUs. For example, an OTU with 1000 reads in the negative control and 3 in a sample, could be removed, but the opposite should not necessarily be done. Examine the context.
- Update training sets for the classification of the sequences and extend them with taxa of interest and perhaps unclassified sequences or sequences with a lower bootstrap

value. Trim the database to the amplified (variable) region(s), since short amplicons could erroneously be matched to a different variable region of a wrong taxon (so called mapping errors).

- Use high bootstrap values (≥ 80 %) when classifying short sequences.

I.3. Data analysis

- Low abundance sequences or OTUs. These don't have any major impacts on estimating β-diversity, although they might be considered to be removed in extensive datasets (e.g., > 10,000 OTUs), since they will slow down computation. They can, however, have a large impact on α-diversity. Erroneous sequences will inflate the perceived diversity. If budget and logistics allow it, and the description of α-diversity is an important aim, samples could be sequenced in a replicated fashion. Only those OTUs or sequences could then be kept which are recovered from, for example, 2 out of 3 replicates.
- And finally ... Don't be afraid of command line interface! Learn R, preferentially combined with a scripting language (Perl or, preferably, Python), a Unix/Linux environment and regular expressions (regex).

Addendum II: Mock community composition

Genus	Species	Reference	Gram stain	Phylum
Arthrobacter	sp.	R-36537	+	*Actinobacteria*
Arthrobacter	sp.	R-36671	+	*Actinobacteria*
Bacillus	sp.	R-43903	+	*Firmicutes*
Brevundimonas	sp.	R-36741	-	*Alphaproteobacteria*
Deinococcus	sp.	R-36593	+	*Deinococcus-Thermus*
Deinococcus	sp.	R-36590	+	*Deinococcus-Thermus*
Deinococcus	sp.	R-36206	+	*Deinococcus-Thermus*
Devosia	sp.	R-43424	-	*Alphaproteobacteria*
Flavobacterium	*aquatile*	LMG4008	-	*Bacteroidetes*
Flavobacterium	*micromati*	R-36963	-	*Bacteroidetes*
Gillisia	sp.	R-39531	-	*Bacteroidetes*
Herbaspirillum	sp.	R-36369	-	*Betaproteobacteria*
Hymenobacter	sp.	R-36591	-	*Bacteroidetes*
Loktanella	*salsilacus*	R-8904	-	*Alphaproteobacteria*
Paenibacillus	*wynnii*	LMG22176	+	*Firmicutes*
Polaromonas	sp.	R-39156	-	*Betaproteobacteria*
Porphyrobacter	sp.	R-39130	-	*Alphaproteobacteria*
Psychroflexus	sp.	R-39535	-	*Bacteroidetes*
Rhodococcus	*fascians*	R-37549	+	*Actinobacteria*
Rothia	sp.	R-36663	+	*Actinobacteria*
Staphylococcus	*warneri*	R-36520	+	*Firmicutes*

References

Abramovich, R.S., Pomati, F., Jungblut, A.D., Guglielmin, M., and Neilan, B.A. (2012) T-RFLP Fingerprinting Analysis of Bacterial Communities in Debris Cones, Northern Victoria Land, Antarctica. *Permafr. Periglac. Process.* **23**: 244–248.

Acinas, S.G., Marcelino, L. A., Klepac-Ceraj, V., and Polz, M.F. (2004) Divergence and redundancy of 16S rRNA sequences in genomes with multiple rrn operons. *J. Bacteriol.* **186**: 2629–2635.

Adams, B.J., Bardgett, R.D., Ayres, E., Wall, D.H., Aislabie, J., Bamforth, S., et al. (2006) Diversity and distribution of Victoria Land biota. *Soil Biol. Biochem.* **38**: 3003–3018.

Aislabie, J.M., Jordan, S., Ayton, J., Klassen, J.L., Barker, G.M., and Turner, S. (2009) Bacterial diversity associated with ornithogenic soil of the Ross Sea region, Antarctica. *Can. J. Microbiol.* **55**: 21–36.

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

Aislabie, J.M., Lau, A., Dsouza, M., Shepherd, C., Rhodes, P., and Turner, S.J. (2013) Bacterial composition of soils of the Lake Wellman area, Darwin Mountains, Antarctica. *Extremophiles* **17**: 775–86.

Alonso-Sáez, L., Zeder, M., Harding, T., Pernthaler, J., Lovejoy, C., Bertilsson, S., and Pedrós-Alió, C. (2014) Winter bloom of a rare betaproteobacterium in the Arctic Ocean. *Front. Microbiol.* **5**: 425.

Amend, A.S., Oliver, T.A., Amaral-Zettler, L.A., Boetius, A., Fuhrman, J.A., Horner-Devine, M.C., et al. (2013) Macroecological patterns of marine bacteria on a global scale. *J. Biogeogr.* **40**: 800–811.

Amend, A.S., Seifert, K. A., Samson, R., and Bruns, T.D. (2010) Indoor fungal composition is geographically patterned and more diverse in temperate zones than in the tropics. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 13748–53.

Andam, C.P., Doroghazi, J.R., Campbell, A.N., Kelly, P.J., Choudoir, M.J., and Buckley, D.H. (2016) A Latitudinal Diversity Gradient in Terrestrial Bacteria of the Genus Streptomyces. *MBio* **7**: 1–9.

Anesio, A.M. and Bellas, C.M. (2011) Are low temperature habitats hot spots of microbial evolution driven by viruses? *Trends Microbiol.* **19**: 52–57.

Anesio, A.M., Hodson, A.J., Fritz, A., Psenner, R., and Sattler, B. (2009) High microbial activity on glaciers: Importance to the global carbon cycle. *Glob. Chang. Biol.* **15**: 955–960.

Archer, S.D.J., McDonald, I.R., Herbold, C.W., Lee, C.K., and Cary, C.S. (2015) Benthic microbial communities of coastal terrestrial and ice shelf Antarctic meltwater ponds. *Front. Microbiol.* **6**: 1–11.

Archer, S.D.J., McDonald, I.R., Herbold, C.W., Lee, C.K., Niederberger, T.S., and Cary, C.

(2015) Temporal, regional and geochemical drivers of microbial community variation in the melt ponds of the Ross Sea region, Antarctica. *Polar Biol.*

Ashworth, A.C. and Kuschel, G. (2003) Fossil weevils (Coleoptera: Curculionidae) from latitude 85°S Antarctica. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **191**: 191–202.

Ashworth, A.C. and Preece, R.C. (2003) The first freshwater molluscs from Antarctica. *J. Molluscan Stud.* **69**: 89–92.

Ashworth, A.C. and Thompson, F.C. (2003) Palaeontology: A fly in the biogeographic ointment. *Nature* **423**: 135–136.

Baker, G.C., Tow, L.A., and Cowan, D. A. (2003) PCR-based detection of non-indigenous microorganisms in "pristine" environments. *J. Microbiol. Methods* **53**: 157–164.

Bakermans, C., Skidmore, M.L., Douglas, S., and McKay, C.P. (2014) Molecular characterization of bacteria from permafrost of the Taylor Valley, Antarctica. *FEMS Microbiol. Ecol.* **89**: 331–346.

Bakermans, C., Tsapin, A.I., Souza-Egipsy, V., Gilichinsky, D. a, and Nealson, K.H. (2003) Reproduction and metabolism at - 10 degrees C of bacteria isolated from Siberian permafrost. *Environ. Microbiol.* **5**: 321–326.

Banks, J.C., Cary, S.C., and Hogg, I.D. (2009) The phylogeography of Adelie penguin faecal flora. *Environ. Microbiol.* **11**: 577–88.

Bargagli, R. (2005) Antarctica: Geomorphology and Climate Trends. In, *Antarctic ecosystems: Environmental contamination, climate change, and human impact*., pp. 1–41.

Barrett, J.E., Virginia, R.A., Hopkins, D.W., Aislabie, J., Bargagli, R., Bockheim, J.G., et al. (2006) Terrestrial ecosystem processes of Victoria Land, Antarctica. *Soil Biol. Biochem.* **38**: 3019–3034.

Barrett, J.E., Virginia, R.A., Wall, D.H., Cary, S.C., Adams, B.J., Hacker, A.L., and Aislabie, J.M. (2006) Co-variation in soil biodiversity and biogeochemistry in northern and southern Victoria Land, Antarctica. *Antarct. Sci.* **18**: 535.

Barrientos-Díaz, L., Gidekel, M., and Gutiérrez-Moraga, A. (2008) Characterization of rhizospheric bacteria isolated from Deschampsia antarctica Desv. *World J. Microbiol. Biotechnol.* **24**: 2289–2296.

Barton, H.A., Taylor, N.M., Kreate, M.P., Springer, A.C., Oehrle, S.A., and Bertog, J.L. (2007) The impact of host rock geochemistry on bacterial community structure in oligotrophic cave environments. *Int. J. Speleol.* **36**: 93–104.

Battista, J.R. (1997) Against all odds: the survival strategies of *Deinococcus radiodurans*. *Annu. Rev. Microbiol.* **51**: 203–224.

Bell, T.H., Yergeau, E., Martineau, C., Juck, D., Whyte, L.G., and Greer, C.W. (2011) Identification of nitrogen-incorporating bacteria in petroleum-contaminated arctic soils by using [15N]DNA-based stable isotope probing and pyrosequencing. *Appl. Environ. Microbiol.* **77**: 4163–4171.

Berry, D., Ben Mahfoudh, K., Wagner, M., and Loy, A. (2012) Barcoded Primers Used in

Multiplex Amplicon Pyrosequencing Bias Amplification. *Appl. Environ. Microbiol.* **78**: 612–612.

Bienhold, C., Boetius, A., and Ramette, A. (2012) The energy–diversity relationship of complex bacterial communities in Arctic deep-sea sediments. *ISME J.* **6**: 724–732.

Bijl, P.K., Bendle, J.A.P., Bohaty, S.M., Pross, J., Schouten, S., Tauxe, L., et al. (2013) Eocene cooling linked to early flow across the Tasmanian Gateway. *Proc. Natl. Acad. Sci. U. S. A.* **110**: 9645–50.

Blazewicz, S.J., Barnard, R.L., Daly, R. A., and Firestone, M.K. (2013) Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *ISME J.* **7**: 2061–8.

Boger, S.D. (2011) Antarctica - Before and after Gondwana. *Gondwana Res.* **19**: 335–371.

Bohorquez, L.C., Delgado-Serrano, L., López, G., Osorio-Forero, C., Klepac-Ceraj, V., Kolter, R., et al. (2012) Erratum to: In-depth Characterization via Complementing Culture-Independent Approaches of the Microbial Community in an Acidic Hot Spring of the Colombian Andes. *Microb. Ecol.* **63**: 238–238.

Borcard, D. and Legendre, P. (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol. Modell.* **153**: 51–68.

Borcard, D., Legendre, P., Avois-Jacquet, C., and Tuomisto, H. (2004) Dissecting the spatial structure of ecological data at multiple scales. *Ecology* **85**: 1826–1832.

Borghini, F., Colacevich, A., and Bargagli, R. (2007) Water geochemistry and sedimentary pigments in northern Victoria Land lakes, Antarctica. *Polar Biol.* **30**: 1173–1182.

Bowman, J.P., Mccammon, S. A., Brown, J.L., and Mcmeekin, T. A. (1998) *Glaciecola punicea* gen. nov., sp. nov. and *Glaciecola pallidula* gen. nov., sp. nov. : psychrophilic bacteria from Antarctic sea-ice habitats. *Int. J. Syst. Bacteriol.* 1213–1222.

Bowman, J.P., McCammon, S. A., Rea, S.M., and McMeekin, T.A. (2000) The microbial composition of three limnologically disparate hypersaline Antarctic lakes. *FEMS Microbiol. Lett.* **183**: 81–8.

Bowman, J.P., Rea, S.M., McCammon, S. A., and McMeekin, T. A. (2000) Diversity and community structure within anoxic sediment from marine salinity meromictic lakes and a coastal meromictic marine basin, Vestfold Hilds, Eastern Antarctica. *Environ. Microbiol.* **2**: 227–237.

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

Brinkmann, M., Pearce, D.A., Convey, P., and Ott, S. (2007) The cyanobacterial community of polygon soils at an inland Antarctic nunatak. *Polar Biol.* **30**: 1505–1511.

Bromwich, D.H., Nicolas, J.P., Monaghan, A.J., Lazzara, M. A., Keller, L.M., Weidner, G. A., and Wilson, A.B. (2012) Central West Antarctica among the most rapidly warming regions on Earth. *Nat. Geosci.* **6**: 139–145.

Brown, M. V, Schwalbach, M.S., Hewson, I., and Fuhrman, J. A. (2005) Coupling 16S-ITS rDNA clone libraries and automated ribosomal intergenic spacer analysis to show marine microbial diversity: development and application to a time series. *Environ. Microbiol.* **7**: 1466–79.

Bryant, D.A., Costas, A.M.G., Maresca, J.A., Chew, A.G.M., Klatt, C.G., Bateson, M.M., et al. (2007) Candidatus *Chloracidobacterium thermophilum*: An Aerobic Phototrophic Acidobacterium. *Science (80-.).* **317**: 523–526.

Burkins, M.B., Virginia, R. A., Chamberlain, C.P., and Wall, D.H. (2011) Origin and Distribution of Soil Organic Matter in Taylor Valley , Antarctica. *Ecology* **81**: 2377–2391.

Cangelosi, G.A. and Meschke, J.S. (2014) Dead or alive: Molecular assessment of microbial viability. *Appl. Environ. Microbiol.* **80**: 5884–5891.

Cannone, N., Ellis-Evans, J.C., Strachan, R., and Guglielmin, M. (2006) Interactions between climate, vegetation and the active layer in soils at two Maritime Antarctic sites. **18**: 323–333.

Cannone, N. and Guglielmin, M. (2009) Influence of vegetation on the ground thermal regime in continental Antarctica. *Geoderma* **151**: 215–223.

Cannone, N., Wagner, D., Hubberten, H.W., and Guglielmin, M. (2008) Biotic and abiotic factors influencing soil properties across a latitudinal gradient in Victoria Land, Antarctica. *Geoderma* **144**: 50–65.

Carpenter, E.J., Lin, S., and Capone, D.G. (2000) Bacterial Activity in South Pole Snow. *Appl. Environ. Microbiol.* **66**: 4514–4517.

Carr, S.A., Orcutt, B.N., Mandernack, K.W., and Spear, J.R. (2015) Abundant *Atribacteria* in deep marine sediment from the Adélie Basin, Antarctica. *Front. Microbiol.* **6**: 1–12.

Cary, S.C., McDonald, I.R., Barrett, J.E., and Cowan, D. A. (2010) On the rocks: the microbiology of Antarctic Dry Valley soils. *Nat. Rev. Microbiol.* **8**: 129–138.

Cavicchioli, R. (2015) Microbial ecology of Antarctic aquatic systems. *Nat. Rev. Microbiol.* **13.**:

Cavicchioli, R., Siddiqui, K.S., Andrews, D., and Sowers, K.R. (2002) Low-temperature extremophiles and their applications. *Curr. Opin. Biotechnol.* **13**: 253–261.

Chakravorty, S., Helb, D., Burday, M., Connell, N., and Alland, D. (2007) A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *J. Microbiol. Methods* **69**: 330–9.

Chan, Y., Lacap, D.C., Lau, M.C.Y., Ha, K.Y., Warren-Rhodes, K. A., Cockell, C.S., et al. (2012) Hypolithic microbial communities: Between a rock and a hard place. *Environ. Microbiol.* **14**: 2272–2282.

Check Hayden, E. (2014) Data from pocket-sized genome sequencer unveiled. *Nature* 2–4.

Cho, J.C. and Tiedje, J.M. (2000) Biogeography and degree of endemicity of fluorescent Pseudomonas strains in soil. *Appl. Environ. Microbiol.* **66**: 5448–5456.

Chong, C.W., Annie Tan, G.Y., Wong, R.C.S., Riddle, M.J., and Tan, I.K.P. (2009) DGGE fingerprinting of bacteria in soils from eight ecologically different sites around Casey Station, Antarctica. *Polar Biol.* **32**: 853–860.

Chong, C.W., Goh, Y.S., Convey, P., Pearce, D. A., and Tan, I.K.P. (2013) Spatial pattern in Antarctica: What can we learn from Antarctic bacterial isolates? *Extremophiles* **17**: 733–745.

Chong, C.W., Pearce, D. A., Convey, P., Yew, W.C., and Tan, I.K.P. (2012) Patterns in the distribution of soil bacterial 16S rRNA gene sequences from different regions of Antarctica. *Geoderma* **181-182**: 45–55.

Chong, C.-W., Pearce, D.A., and Convey, P. (2015) Emerging spatial patterns in Antarctic prokaryotes. *Front. Microbiol.* **6.**:

Chown, S.L. and Convey, P. (2007) Spatial and temporal variability across life's hierarchies in the terrestrial Antarctic. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **362**: 2307–2331.

Christner, B.C., Priscu, J.C., Achberger, A.M., Barbante, C., Carter, S.P., Christianson, K., et al. (2014) A microbial ecosystem beneath the West Antarctic ice sheet. *Nature* **512**: 310–313.

Clarke, K.R., Somerfield, P.J., and Gorley, R.N. (2008) Testing of null hypotheses in exploratory community analyses: similarity profiles and biota-environment linkage. *J. Exp. Mar. Bio. Ecol.* **366**: 56–69.

Cleenwerck, I., Camu, N., Engelbeen, K., De Winter, T., Vandemeulebroecke, K., De Vos, P., and De Vuyst, L. (2007) *Acetobacter ghanensis* sp. nov., a novel acetic acid bacterium isolated from traditional heap fermentations of Ghanaian cocoa beans. *Int. J. Syst. Evol. Microbiol.* **57**: 1647–1652.

Convey, P. (1996) The influence of environmental characteristics on life history attributes of Antarctic terrestrial biota. *Biol. Rev.* **71**: 191–225.

Convey, P. and Block, W. (1996) Antarctic Diptera: ecology, physiology and distribution. *Eur. J. Entomol.* **93**: 1–13.

Convey, P., Chown, S.L., Clarke, A., Barnes, D.K. A., Bokhorst, S., Cummings, V., et al. (2014) The spatial structure of Antarctic biodiversity. *Ecol. Monogr.* **84**: 203–244.

Convey, P., Gibson, J. A. E., Hillenbrand, C.-D., Hodgson, D. A., Pugh, P.J. A., Smellie, J.L., and Stevens, M.I. (2008) Antarctic terrestrial life--challenging the history of the frozen continent? *Biol. Rev. Camb. Philos. Soc.* **83**: 103–117.

Convey, P. and McInnes, S.J. (2005) Exceptional tardigrade-dominated ecosystems in Ellsworth Land, Antarctica. *Ecology* **86**: 519–527.

Convey, P. and Smith, R.I.L. (2006) Responses of terrestrial Antarctic ecosystems to climate change. *Plant Ecol.* **182**: 1–10.

Convey, P. and Stevens, M.I. (2007) Ecology. Antarctic biodiversity. *Science* **317**: 1877–1878.

Corinaldesi, C., Danovaro, R., and Dell'Anno, A. (2005) Simultaneous Recovery of Extracellular and Intracellular DNA Suitable for Molecular Studies from Marine

Sediments Simultaneous Recovery of Extracellular and Intracellular DNA Suitable for Molecular Studies from Marine Sediments. *Appl. Environ. Microbiol.* **71**: 46–50.

Cowan, D. A., Khan, N., Pointing, S.B., and Cary, S.C. (2010) Diverse hypolithic refuge communities in the McMurdo Dry Valleys. *Antarct. Sci.* **22**: 714–720.

Cowan, D.A. and Ah Tow, L. (2004) Endangered Antarctic environments. *Annu. Rev. Microbiol.* **58**: 649–690.

Cowan, D.A., Chown, S.L., Convey, P., Tuffin, M., Hughes, K., Pointing, S., and Vincent, W.F. (2011) Non-indigenous microorganisms in the Antarctic: Assessing the risks. *Trends Microbiol.* **19**: 540–548.

Cowan, D.A., Makhalanyane, T.P., Dennis, P.G., and Hopkins, D.W. (2014) Microbial ecology and biogeochemistry of continental Antarctic soils. *Front. Microbiol.* **5**: 154.

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

Cowan, D.A., Sohm, J.A., Makhalanyane, T.P., Capone, D.G., Green, T.G.. A., Cary, S.C., and Tuffin, I.M. (2011) Hypolithic communities: Important nitrogen sources in Antarctic desert soils. *Environ. Microbiol. Rep.* **3**: 581–586.

D'Amico, S., Collins, T., Marx, J.-C., Feller, G., and Gerday, C. (2006) Psychrophilic microorganisms: challenges for life. *EMBO Rep.* **7**: 385–389.

Dana, G.L., Wharton Jr., R.A., and Dubayah, R.A. (1998) Solar radiation in the Mcmurdo Dry Valleys, Antarctica. In, *Ecosystem Dynamics in a Polar Desert: The McMurdo Dry Valleys, Antarctica*., pp. 39–64.

Das, D., Baruah, R., Sarma Roy, A., Singh, A.K., Deka Boruah, H.P., Kalita, J., and Bora, T.C. (2015) Complete genome sequence analysis of Pseudomonas aeruginosa N002 reveals its genetic adaptation for crude oil degradation. *Genomics* **105**: 182–190.

Datta, S., Datta, S., Kim, S., Chakraborty, S., and Gill, R.S. (2010) Statistical Analyses of Next Generation Sequence Data: A Partial Overview. *J. Proteomics Bioinform.* **03**: 183–190.

Davis, K.E.R., Joseph, S.J., and Janssen, P.H. (2005) Effects of growth medium, inoculum size, and incubation time on culturability and isolation of soil bacteria. *Appl. Environ. Microbiol.* **71**: 826–834.

Day, T.A., Ruhland, C.T., Grobe, C.W., and Xiong, F. (1999) Growth and reproduction of Antarctic vascular plants in response to warming and UV radiation reductions in the field. *Oecologia* **119**: 24–35.

Delavat, F., Lett, M.-C., and Lievremont, D. (2012) Novel and unexpected bacterial diversity in an arsenic-rich ecosystem revealed by culture-dependent approaches. *Biol. Direct* **7**: 28.

DeSantis, T.Z., Hugenholtz, P., Keller, K., Brodie, E.L., Larsen, N., Piceno, Y.M., et al. (2006) NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic Acids Res.* **34**: W394–9.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., et al. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**: 5069–5072.

Devos, D.P. (2013) *Gemmata obscuriglobus*. *Curr. Biol.* **23**: R705–R707.

Dickinson, W.W. and Rosen, M.R. (2003) Antarctic permafrost: An analogue for water and diagenetic minerals on Mars. *Geology* **31**: 199–202.

Dickson, J.L., Head, J.W., Levy, J.S., and Marchant, D.R. (2013) Don Juan Pond, Antarctica: Near-surface CaCl(2)-brine feeding Earth's most saline lake and implications for Mars. *Sci. Rep.* **3**: 1166.

Donachie, S.P., Foster, J.S., and Brown, M. V (2007) Culture clash: challenging the dogma of microbial diversity. *ISME J.* **1**: 97–99.

van Dorst, J., Bissett, A., Palmer, A.S., Brown, M., Snape, I., Stark, J.S., et al. (2014) Community fingerprinting in a sequencing world. *FEMS Microbiol. Ecol.* **89**: 316–330.

Dray, S., Legendre, P., and Peres-Neto, P.R. (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecol. Modell.* **196**: 483–493.

Ducklow, H.W., Fraser, W., Meredith, M., Stammerjohn, S., Doney, S.C., Martinson, D., et al. (2013) West Antarctic Peninsula: An Ice-Dependent Coastal Marine Ecosystem in Transition. *Oceanography* **26**: 190–203.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–1.

Edgar, R.C. (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10**: 996–8.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194–200.

Edwards, U., Rogall, T., Blöcker, H., Emde, M., and Böttger, E.C. (1989) Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. *Nucleic Acids Res.* **17**: 7843–7853.

Ellis-Evans, J.C. (1996) Microbial diversity and function in Antarctic freshwater ecosystems. *Biodivers. Conserv.* **5**: 1395–1431.

Engelbrektson, A., Kunin, V., Wrighton, K.C., Zvenigorodsky, N., Chen, F., Ochman, H., and Hugenholtz, P. (2010) Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *ISME J.* **4**: 642–647.

Eren, A.M., Maignien, L., Sul, W.J., Murphy, L.G., Grim, S.L., Morrison, H.G., and Sogin, M.L. (2013) Oligotyping: Differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol. Evol.* **4**: 1111–1119.

Eren, A.M., Morrison, H.G., Lescault, P.J., Reveillaud, J., Vineis, J.H., and Sogin, M.L. (2015) Minimum entropy decomposition: unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *Isme J* **9**: 968–979.

Ertz, D., Aptroot, A., Van de Vijver, B., Sliwa, L., Moermans, C., and Øvstedal, D. (2014) Lichens from the Utsteinen Nunatak (Sør Rondane Mountains, Antarctica), with the description of one new species and the establishment of permanent plots. *Phytotaxa* **191**: 99.

Faith, J.J., Guruge, J.L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A.L., et al. (2013) The long-term stability of the human gut microbiota. *Science (80-.).* **341**: 1237439.

Felske, A. and Akkermans, A.D. (1998) Prominent occurrence of ribosomes from an uncultured bacterium of the Verrucomicrobiales cluster in grassland soils. *Lett. Appl. Microbiol.* **26**: 219–23.

Fenchel, T. and Finlay, B.J. (2004) The Ubiquity of Small Species: Patterns of Local and Global Diversity. *Bioscience* **54**: 777.

Feng, Y., Grogan, P., Caporaso, J.G., Zhang, H., Lin, X., Knight, R., and Chu, H. (2014) pH is a good predictor of the distribution of anoxygenic purple phototrophic bacteria in Arctic soils. *Soil Biol. Biochem.* **74**: 193–200.

Fernandez-Carazo, R., Hodgson, D.A., Convey, P., and Wilmotte, A. (2011) Low cyanobacterial diversity in biotopes of the Transantarctic Mountains and Shackleton Range (80-82° S), Antarctica. *FEMS Microbiol. Ecol.* **77**: 503–517.

Fernández-Carazo, R., Namsaraev, Z., Mano, M.-J., Ertz, D., and Wilmotte, A. (2012) Cyanobacterial diversity for an anthropogenic impact assessment in the Sør Rondane Mountains area, Antarctica. *Antarct. Sci.* **24**: 229–242.

Fierer, N., Bradford, M. A., and Jackson, R.B. (2007) Toward an ecological classification of soil bacteria. *Ecology* **88**: 1354–64.

Fierer, N. and Jackson, R.B. (2006) The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 626–631.

Fine, P. (2015) Ecological and Evolutionary Drivers of Geographic Variation in Species Diversity. *Annu. Rev. Ecol. Evol. Syst.* **46**: 369–392.

Fisher, M.M. and Triplett, E.W. (1999) Automated Approach for Ribosomal Intergenic Spacer Analysis of Microbial Diversity and Its Application to Freshwater Bacterial Communities. **65**: 4630–4636.

Fontaneto, D., Iakovenko, N., and De Smet, W.H. (2015) Diversity gradients of rotifer species richness in Antarctica. *Hydrobiologia* **761**: 235–248.

Franzetti, A., Tatangelo, V., Gandolfi, I., Bertolini, V., Bestetti, G., Diolaiuti, G., et al. (2013) Bacterial community structure on two alpine debris-covered glaciers and biogeography of *Polaromonas* phylotypes. 1483–1492.

Franzmann, P.D. (1996) Examination of Antarctic prokaryotic diversity through molecular comparisons. *Biodivers. Conserv.* **5**: 1295–1305.

Franzmann, P.D., Burton, H.R., and Mcmeekin, T. A. (1987) *Halomonas subglaciescola*, a New Species of Halotolerant Bacteria Isolated from Antarctica. *Int. J. Syst. Bacteriol.* **37**: 27–34.

Franzmann, P.D. and Dobson, S.J. (1993) The phylogeny of bacteria from a modern Antarctic refuge. *Antarct. Sci.* **5**: 267–270.

Freckman, D.W. and Virginia, R. A. (1997) Low-diversity antarctic soil nematode communities: Distribution and response to disturbance. *Ecology* **78**: 363–369.

Fuhrman, J.A., Steele, J.A., Hewson, I., Schwalbach, M.S., Brown, M. V., Green, J.L., and Brown, J.H. (2008) A latitudinal diversity gradient in planktonic marine bacteria. *Proc. Natl. Acad. Sci.* **105**: 7774–8.

Geyer, K.M., Altrichter, A.E., Van Horn, D.J., Takacs-Vesbach, C.D., Gooseff, M.N., and Barrett, J.E. (2013) Environmental controls over bacterial communities in polar desert soils. *Ecosphere* **4**: art127.

Geyer, K.M., Altrichter, A.E., Takacs-Vesbach, C.D., Van Horn, D.J., Gooseff, M.N., and Barrett, J.E. (2014) Bacterial community composition of divergent soil habitats in a polar desert. *FEMS Microbiol. Ecol.* **89**: 490–494.

Ghiglione, J.-F., Galand, P.E., Pommier, T., Pedros-Alio, C., Maas, E.W., Bakker, K., et al. (2012) Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proc. Natl. Acad. Sci.* **109**: 17633–17638.

Ghiglione, J.F. and Murray, A.E. (2012) Pronounced summer to winter differences and higher wintertime richness in coastal Antarctic marine bacterioplankton. **14**: 617–629.

Gilles, A., Meglécz, E., Pech, N., Ferreira, S., Malausa, T., and Martin, J.-F. (2011) Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Genomics* **12**: 245.

Giovannoni, S. and Stingl, U. (2007) The importance of culturing bacterioplankton in the "omics" age. *Nat. Rev. Microbiol.* **5**: 820–826.

Gomez-Alvarez, V., Teal, T.K., and Schmidt, T.M. (2009) Systematic artifacts in metagenomes from complex microbial communities. *ISME J.* **3**: 1314–7.

Gorodetskaya, I. V., Kneifel, S., Maahn, M., Van Tricht, K., Thiery, W., Schween, J.H., et al. (2015) Cloud and precipitation properties from ground-based remote sensing instruments in East Antarctica. *Cryosph.* **9**: 285–304.

Gorodetskaya, I. V., Van Lipzig, N.P.M., Van Den Broeke, M.R., Mangold, A., Boot, W., and Reijmer, C.H. (2013) Meteorological regimes and accumulation patterns at Utsteinen, Dronning Maud Land, East Antarctica: Analysis of two contrasting years. *J. Geophys. Res. Atmos.* **118**: 1700–1715.

Gottel, N.R., Castro, H.F., Kerley, M., Yang, Z., Pelletier, D. A., Podar, M., et al. (2011) Distinct microbial communities within the endosphere and rhizosphere of *Populus deltoides* roots across contrasting soil types. *Appl. Environ. Microbiol.* **77**: 5934–44.

Grant, A. and Ogilvie, L.A. (2004) Name that microbe: rapid identification of taxa responsible for individual fragments in fingerprints of microbial community structure. *Mol. Ecol. Notes* **4**: 133–136.

Green, J.L. (2014) Microbial Biogeography : From Taxonomy to Traits. **1039.**:

Grégoire, P., Bohli, M., Cayol, J.L., Joseph, M., Guasco, S., Dubourg, K., et al. (2011) *Caldilinea tarbellica* sp. nov., a filamentous, thermophilic, anaerobic bacterium isolated from a deep hot aquifer in the Aquitaine Basin. *Int. J. Syst. Evol. Microbiol.* **61**: 1436–1441.

Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M., and Whiteley, A.S. (2011) The bacterial biogeography of British soils. *Environ. Microbiol.* **13**: 1642–1654.

Grzymski, J.J., Riesenfeld, C.S., Williams, T.J., Dussaq, A.M., Ducklow, H., Erickson, M., et al. (2012) A metagenomic assessment of winter and summer bacterioplankton from Antarctica Peninsula coastal surface waters. *ISME J.* **6**: 1901–1915.

Van der Gucht, K., Cottenie, K., Muylaert, K., Vloemans, N., Cousin, S., Declerck, S., et al. (2007) The power of species sorting: local factors drive bacterial community composition over a wide range of spatial scales. *Proc. Natl. Acad. Sci. U. S. A.* **104**: 20404–20409.

Guerrero, R. and Berlanga, M. (2013) An integrate ecogenetic study of minimal ecosystems: The microbial mats of Ebro Delta and the Camargue (Western Mediterranean). *Contrib. to Sci.* **9**: 117–139.

Gürtler, V. and Stanisich, V. A. (1996) New approaches to typing and identification of bacteria using the 16S-23S rDNA spacer region. *Microbiology* **142**: 3–16.

Haas, B.J., Gevers, D., Earl, A.M., Feldgarden, M., Ward, D. V, Giannoukos, G., et al. (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res.* **21**: 494–504.

Hall, B.L., Denton, G.H., Fountain, A.G., Hendy, C.H., and Henderson, G.M. (2010) Antarctic lakes suggest millennial reorganizations of Southern Hemisphere atmospheric and oceanic circulation. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 21355–21359.

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M.C., and Martiny, J.B.H. (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.* **10**: 1–10.

Harismendy, O., Ng, P.C., Strausberg, R.L., Wang, X., Stockwell, T.B., Beeson, K.Y., et al. (2009) Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biol.* **10**: R32.

He, Y., Caporaso, J.G., Jiang, X.-T., Sheng, H.-F., Huse, S.M., Rideout, J.R., et al. (2015) Stability of operational taxonomic units: an important but neglected property for analyzing microbial diversity. *Microbiome* **3**: 20.

Hirsch, P., Gallikowski, C. A., Siebert, J., Peissl, K., Kroppenstedt, R., Schumann, P., et al. (2004) *Deinococcus frigens* sp. nov., *Deinococcus saxicola* sp. nov., and *Deinococcus marmoris* sp. nov., low temperature and draught-tolerating, UV-resistant bacteria from continental Antarctica. *Syst. Appl. Microbiol.* **27**: 636–645.

Hodgson, D. A. (2011) First synchronous retreat of ice shelves marks a new phase of polar deglaciation. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 18859–60.

Hodgson, D. A., Convey, P., Verleyen, E., Vyverman, W., McInnes, S.J., Sands, C.J., et al. (2010) The limnology and biology of the Dufek Massif, Transantarctic Mountains 82°

South. *Polar Sci.* **4**: 197–214.

Hodgson, D. A., Vyverman, W., and Sabbe, K. (2001) Limnology and biology of saline lakes in the Rauer Islands, eastern Antarctica. *Antarct. Sci.* **13**: 255–270.

Hogg, I.D., Craig Cary, S., Convey, P., Newsham, K.K., O'Donnell, A.G., Adams, B.J., et al. (2006) Biotic interactions in Antarctic terrestrial ecosystems: Are they a factor? *Soil Biol. Biochem.* **38**: 3035–3040.

Van Horn, D.J., Van Horn, M.L., Barrett, J.E., Gooseff, M.N., Altrichter, A.E., Geyer, K.M., et al. (2013) Factors Controlling Soil Microbial Biomass and Bacterial Diversity and Community Composition in a Cold Desert Ecosystem: Role of Geographic Scale. *PLoS One* **8.**:

Howard-Williams, C. and Hawes, I. (2007) Ecological processes in Antarctic inland waters: interactions between physical processes and the nitrogen cycle. *Antarct. Sci.* **19**: 205.

Howard-Williams, C., Hawes, I., and Gordon, S. (2010) The environmental basis of ecosystem variability in Antarctica: Research in the Latitudinal Gradient Project. *Antarct. Sci.* **22**: 591–602.

Huang, J.P., Swain, A.K., Andersen, D.T., and Bej, A.K. (2014) Bacterial diversity within five unexplored freshwater lakes interconnected by surface channels in East Antarctic Dronning Maud Land (Schirmacher Oasis) using amplicon pyrosequencing. 359–366.

Huber, J. A., Mark Welch, D.B., Morrison, H.G., Huse, S.M., Neal, P.R., Butterfield, D. A. and Sogin, M.L. (2007) Microbial population structures in the deep marine biosphere. *Science* **318**: 97–100.

Hugenholtz, P. (2002) Exploring prokaryotic diversity in the genomic era. *Genome Biol.* **3**: REVIEWS0003.

Huse, S.M., Huber, J. A., Morrison, H.G., Sogin, M.L., and Welch, D.M. (2007) Accuracy and quality of massively-parallel DNA pyrosequencing. *Genome Biol.* **8**: R143.

Huse, S.M., Welch, D.M., Morrison, H.G., and Sogin, M.L. (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environ. Microbiol.* **12**: 1889–1898.

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J.H., Chinwalla, A.T., et al. (2012) Structure, function and diversity of the healthy human microbiome. *Nature* **486**: 207–214.

Huybers, P. and Denton, G. (2008) Antarctic temperature at orbital timescales controlled by local summer duration. *Nat. Geosci.* **1**: 787–792.

Jain, M., Fiddes, I.T., Miga, K.H., Olsen, H.E., Paten, B., and Akeson, M. (2015) Improved data analysis for the MinION nanopore sequencer. *Nat. Methods* **12**: 351–356.

Janssen, P.H. (2006) Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl. Environ. Microbiol.* **72**: 1719–1728.

Jeraldo, P., Chia, N., and Goldenfeld, N. (2011) On the suitability of short reads of 16S rRNA for phylogeny-based analyses in environmental surveys. *Environ. Microbiol.* **13**: 3000–9.

Jones, M.B., Highlander, S.K., Anderson, E.L., Li, W., Dayrit, M., Klitgord, N., et al. (2015) Library preparation methodology can influence genomic and functional predictions in human microbiome research. *Proc Natl Acad Sci U S A* **112**: 1519288112–.

Jones, P.D., Lister, D.H., Osborn, T.J., Harpham, C., Salmon, M., and Morice, C.P. (2012) Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010. *J. Geophys. Res. Atmos.* **117.**:

Jones, R.T., Robeson, M.S., Lauber, C.L., Hamady, M., Knight, R., and Fierer, N. (2009) A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *Isme J.* **3**: 442–453.

Kan, J., Wang, K., and Chen, F. (2006) Temporal variation and detection limit of an estuarine bacterioplankton community analyzed by denaturing gradient gel electrophoresis (DGGE). *Aquat. Microb. Ecol.* **42**: 7–18.

Kautz, S., Rubin, B.E.R., Russell, J. a, and Moreau, C.S. (2013) Surveying the microbiome of ants: comparing 454 pyrosequencing with traditional methods to uncover bacterial diversity. *Appl. Environ. Microbiol.* **79**: 525–34.

Kennedy, A.D. (1995) Antarctic terrestrial ecosystem response to global environmental change. *Fautin, D. G. (Ed.). Annu. Rev. Ecol. Syst. Vol. 26. Xiv+757p. Annu. Rev. Inc. Palo Alto, California, Usa.* **26**: 683–704.

Kennedy, A.D. (1993) Water as a limiting factor in the Antarctic Terrestrial Environment: A Biogeographical Synthesis. *Arct. Alp. Res.* **25**: 308–315.

Kennedy, K., Hall, M.W., Lynch, M.D.J., Moreno-hagelsieb, G., and Neufeld, D. (2014) Evaluating Bias of Illumina-Based Bacterial 16S rRNA Gene Profiles. **80**: 5717–5722.

Kim, M., Cho, A., Lim, H.S., Hong, S.G., Kim, J.H., Lee, J., et al. (2015) Highly Heterogeneous Soil Bacterial Communities around Terra Nova Bay of Northern Victoria Land, Antarctica. *PLoS One* **10**: e0119966.

Kim, S.J., Shin, S.C., Hong, S.G., Lee, Y.M., Choi, I.G., and Park, H. (2012) Genome sequence of a novel member of the genus psychrobacter isolated from antarctic soil. *J. Bacteriol.* **194**: 2403–2403.

Kirchman, D., Moran, X., and Ducklow, H. (2009) Microbial growth in the polar oceans — role of temperature and potential impact of climate change. *Nat. Rev. Microbiol.* **7**: 451–459.

Kirchman, D.L., Cottrell, M.T., and Lovejoy, C. (2010) The structure of bacterial communities in the western Arctic Ocean as revealed by pyrosequencing of 16S rRNA genes. *Environ. Microbiol.* **12**: 1132–43.

Knight, R., Jansson, J., Field, D., Fierer, N., Desai, N., Fuhrman, J. A., et al. (2012) Unlocking the potential of metagenomics through replicated experimental design. *Nat. Biotechnol.* **30**: 513–20.

Koeppel, A.F. and Wu, M. (2013) Surprisingly extensive mixed phylogenetic and ecological signals among bacterial Operational Taxonomic Units. *Nucleic Acids Res.* **41**: 5175–5188.

Kovacs, A., Yacoby, K., and Gophna, U. (2010) A systematic assessment of automated ribosomal intergenic spacer analysis (ARISA) as a tool for estimating bacterial richness. *Res. Microbiol.* **161**: 192–7.

Kunin, V., Engelbrektson, A., Ochman, H., and Hugenholtz, P. (2010) Wrinkles in the rare biosphere: Pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ. Microbiol.* **12**: 118–123.

Lagier, J.C., Armougom, F., Million, M., Hugon, P., Pagnier, I., Robert, C., et al. (2012) Microbial culturomics: Paradigm shift in the human gut microbiome study. *Clin. Microbiol. Infect.* **18**: 1185–1193.

Lamprinou, V., Skaraki, K., Kotoulas, G., Economou-Amilli, A., and Pantazidou, A. (2012) *Toxopsis calypsus* gen. nov., sp. nov. (*Cyanobacteria*, *Nostocales*) from cave "Francthi", Peloponnese, Greece: A morphological and molecular evaluation. *Int. J. Syst. Evol. Microbiol.* **62**: 2870–2877.

Landone Vescovo, I.A., Golemba, M.D., Di Lello, F.A., Culasso, A.C.A., Levin, G., Ruberto, L., et al. (2014) Rich bacterial assemblages from Maritime Antarctica (Potter Cove, South Shetlands) reveal several kinds of endemic and undescribed phylotypes. *Rev. Argent. Microbiol.* **46**: 218–230.

Lauro, F.M., DeMaere, M.Z., Yau, S., Brown, M. V, Ng, C., Wilkins, D., et al. (2011) An integrative study of a meromictic lake ecosystem in Antarctica. *ISME J.* **5**: 879–95.

Laver, T., Harrison, J., O'Neill, P.A., Moore, K., Farbos, A., Paszkiewicz, K., and Studholme, D.J. (2015) Assessing the performance of the Oxford Nanopore Technologies MinION. *Biomol. Detect. Quantif.* **3**: 1–8.

Lawley, B., Ripley, S., Bridge, P., and Convey, P. (2004) Molecular analysis of geographic patterns of eukaryotic diversity in antarctic soils. *Appl. Environ. Microbiol.* **70**: 5963–5972.

Laybourn-Parry, J. (2002) Survival mechanisms in Antarctic lakes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **357**: 863–9.

Laybourn-Parry, J., Anesio, A.M., Madan, N., and Säwström, C. (2013) Virus dynamics in a large epishelf lake (Beaver Lake, Antarctica). *Freshw. Biol.* **58**: 1484–1493.

Laybourn-Parry, J. and Bell, E.M. (2014) Ace Lake: three decades of research on a meromictic, Antarctic lake. *Polar Biol.* **37**: 1685–1699.

Laybourn-Parry, J., Hofer, J.S., and Sommaruga, R. (2001) Viruses in the plankton of freshwater and saline Antarctic lakes. *Freshw. Biol.* **46**: 1279–1287.

Laybourn-Parry, J. and Pearce, D.A. (2007) The biodiversity and ecology of Antarctic lakes: models for evolution. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **362**: 2273–2289.

Laybourn-Parry, J., Quayle, W.C., Henshaw, T., Ruddell, A., and Marchant, H.J. (2001) Life on the edge: the plankton and chemistry of Beaver Lake, an ultra-oligotrophic epishelf lake, Antarctica. *Freshw. Biol.* **46**: 1205–1217.

Laybourn-Parry, J. and Wadham, J.L. (2014) Antarctic Lakes 1st ed. Oxford.

Lee, C.K., Barbier, B.A., Bottos, E.M., McDonald, I.R., and Cary, S.C. (2012) The Inter-Valley Soil Comparative Survey: the ecology of Dry Valley edaphic microbial communities. *ISME J.* **6**: 1046–1057.

Lee, C.K., Herbold, C.W., Polson, S.W., Wommack, K.E., Williamson, S.J., McDonald, I.R., and Cary, S.C. (2012) Groundtruthing Next-Gen Sequencing for Microbial Ecology-Biases and Errors in Community Structure Estimates from PCR Amplicon Pyrosequencing. *PLoS One* **7**: e44224.

Lee, K.C.Y., Herbold, C.W., Dunfield, P.F., Morgan, X.C., McDonald, I.R., and Stott, M.B. (2013) Phylogenetic delineation of the novel phylum *Armatimonadetes* (former candidate division OP10) and definition of two novel candidate divisions. *Appl. Environ. Microbiol.* **79**: 2484–7.

Leff, J.W., Nemergut, D.R., Grandy, A. S., O'Neill, S.P., Wickings, K., Townsend, A.R., and Cleveland, C.C. (2011) The Effects of Soil Bacterial Community Structure on Decomposition in a Tropical Rain Forest. *Ecosystems* **15**: 284–298.

Legendre, P., Borcard, D., and Peres-Neto, P.R. (2005) Analyzing beta diversity: Partitioning the spatial variation of community composition data. *Ecol. Monogr.* **75**: 435–450.

Liu, W., Marsh, T.L., and Cheng, H. (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S Characterization of Microbial Diversity by Determining Terminal Restriction Fragment Length Polymorphisms of Genes Encoding 16S.

Lomolino, M. V, Riddle, B.R., and Brown, J.H. (2006) Biogeography Third Edit. Sinauer,A.D. (ed) Sinauer Associates Inc.

Lopatina, A., Krylenkov, V., and Severinov, K. (2013) Activity and bacterial diversity of snow around russian antarctic stations. *Res. Microbiol.* **164**: 949–958.

Loperena, L., Soria, V., Varela, H., Lupo, S., Bergalli, A., Guigou, M., et al. (2012) Extracellular enzymes produced by microorganisms isolated from maritime Antarctica. *World J. Microbiol. Biotechnol.* **28**: 2249–2256.

López-Bueno, A., Tamames, J., Velázquez, D., Moya, A., Quesada, A., and Alcamí, A. (2009) High diversity of the viral community from an Antarctic lake. *Science* **326**: 858–61.

de los Ríos, A., Ascaso, C., Wierzchos, J., Vincent, W.F., and Quesada, A. (2015) Microstructure and cyanobacterial composition of microbial mats from the High Arctic. *Biodivers. Conserv.* **24**: 476–482.

Lundin, D., Severin, I., Logue, J.B., Ostman, O., Andersson, A.F., and Lindström, E.S. (2012) Which sequencing depth is sufficient to describe patterns in bacterial α- and β-diversity? *Environ. Microbiol. Rep.* **4**: 367–72.

Luo, C., Tsementzi, D., Kyrpides, N., Read, T., and Konstantinidis, K.T. (2012) Direct Comparisons of Illumina vs. Roche 454 Sequencing Technologies on the Same Microbial Community DNA Sample. *PLoS One* **7**: e30087.

Lythe, M., Vaughan, D.G., Lambrecht, A., Miller, H., Nixdorf, U., Oerter, H., et al. (2001) BEDMAP: a new ice thickness and subglacial topographic model of Antarctica. *J.*

Geophys. Res. **106**: 11335–11351.

Maestre, F.T., Delgado-Baquerizo, M., Jeffries, T.C., Eldridge, D.J., Ochoa, V., Gozalo, B., et al. (2015) Increasing aridity reduces soil microbial diversity and abundance in global drylands. *Proc. Natl. Acad. Sci. U. S. A.* **112**: 15684–15689.

Magalhães, C., Stevens, M.I., Cary, S.C., Ball, B. A., Storey, B.C., Wall, D.H., et al. (2012) At limits of life: multidisciplinary insights reveal environmental constraints on biotic diversity in continental Antarctica. *PLoS One* **7**: e44578.

Maitra, R.D., Kim, J., and Dunbar, W.B. (2012) Recent advances in nanopore sequencing. *Electrophoresis* **33**: 3418–3428.

Makhalanyane, T.P., Valverde, A., Birkeland, N.-K., Cary, S.C., Tuffin, I.M., and Cowan, D. a (2013) Evidence for successional development in Antarctic hypolithic bacterial communities. *ISME J.* **7**: 2080–2090.

Margesin, R., Dieplinger, H., Hofmann, J., Sarg, B., and Lindner, H. (2005) A cold-active extracellular metalloprotease from Pedobacter cryoconitis - Production and properties. *Res. Microbiol.* **156**: 499–505.

Margesin, R. and Miteva, V. (2011) Diversity and ecology of psychrophilic microorganisms. *Res. Microbiol.* **162**: 346–361.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L. A., et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**: 376–80.

Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J. A., Green, J.L., et al. (2006) Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol.* **4**: 102–112.

Matsuoka, N., Thomachot, C.E., Oguchi, C.T., Hatta, T., Abe, M., and Matsuzaki, H. (2006) Quaternary bedrock erosion and landscape evolution in the Sør Rondane Mountains, East Antarctica: Reevaluating rates and processes. *Geomorphology* **81**: 408–420.

McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., et al. (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **6**: 610–618.

McKenzie, R.L., Aucamp, P.J., Bais, A.F., Björn, L.O., Ilyas, M., and Madronich, S. (2011) Ozone depletion and climate change: impacts on UV radiation. *Photochem. Photobiol. Sci.* **10**: 182–198.

McKnight, D.M., Runkel, R.L., Tate, C.M., Duff, J.H., and Moorhead, D.L. (2004) Inorganic N and P dynamics of Antarctic glacial meltwater streams as controlled by hyporheic exchange and benthic autotrophic communities. *J. North Am. Benthol. Soc.* **23**: 171–188.

Meola, M., Lazzaro, A., and Zeyer, J. (2014) Diversity, resistance and resilience of the bacterial communities at two alpine glacier forefields after a reciprocal soil transplantation. *Environ. Microbiol.* **16**: 1918–1934.

Michaud, L., Caruso, C., Mangano, S., Interdonato, F., Bruni, V., and Lo Giudice, A. (2012) Predominance of *Flavobacterium*, *Pseudomonas*, and *Polaromonas* within the

prokaryotic community of freshwater shallow lakes in the northern Victoria Land, East Antarctica. *FEMS Microbiol. Ecol.* **82**: 391–404.

Mikheyev, A.S. and Tin, M.M.Y. (2014) A first look at the Oxford Nanopore MinION sequencer. *Mol. Ecol. Resour.* **14**: 1097–1102.

Mikucki, J. a. and Priscu, J.C. (2007) Bacterial diversity associated with blood falls, a subglacial outflow from the Taylor Glacier, Antarctica. *Appl. Environ. Microbiol.* **73**: 4029–4039.

Miller, J.R., Koren, S., and Sutton, G. (2010) Assembly algorithms for next-generation sequencing data. *Genomics* **95**: 315–27.

Misra, H.S., Rajpurohit, Y.S., and Kota, S. (2013) Physiological and molecular basis of extreme radioresistance in *Deinococcus radiodurans*. *Curr. Sci.* **104**: 194–205.

Mitchell, A.C., Lafrenière, M.J., Skidmore, M.L., and Boyd, E.S. (2013) Influence of bedrock mineral composition on microbial diversity in a subglacial environment. *Geology* **41**: 855–858.

Monaghan, A.J., Bromwich, D.H., Fogt, R.L., Wang, S.-H., Mayewski, P.A., Dixon, D.A., et al. (2006) Insignificant change in Antarctic snowfall since the International Geophysical Year. *Science (80-.).* **313**: 827–831.

Moro, D., Lawson, M. A., Hobbs, R.P., and Thompson, R.C. A. (2003) Pathogens of house mice on arid Boullanger Island and subantarctic Macquarie Island, Australia. *J. Wildl. Dis.* **39**: 762–71.

Mueller, D.R. and Vincent, W.F. (2003) Break-up of the largest Arctic ice shelf and associated loss of an epishelf lake. *Geophys. Res. Lett.* **30**: 1–4.

Muyzer, G., De Waal, E.C., and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by DGGE analysis of PCR amplified genes coding for 16S rRNA. *Appl Env. Microbiol* **59**: 695–700.

Naish, T.R., Powell, R.D., Levy, R., Wilson, G.S., Scherer, R.P., Talarico, F., et al. (2009) Obliquity-paced Pliocene West Antarctic ice sheet oscillations. *Nature* **458**: 322–328.

Namsaraev, Z., Mano, M.-J., Fernandez, R., and Wilmotte, A. (2010) Biogeography of terrestrial cyanobacteria from Antarctic ice-free areas. *Ann. Glaciol.* **51**: 171–177.

Newsham, K.K., Pearce, D. A., and Bridge, P.D. (2010) Minimal influence of water and nutrient content on the bacterial community composition of a maritime Antarctic soil. *Microbiol. Res.* **165**: 523–530.

Newton, I.L. and Roeselers, G. (2012) The effect of training set on the classification of honey bee gut microbiota using the Naive Bayesian Classifier. *BMC Microbiol.* **12**: 221.

Ng, C., DeMaere, M.Z., Williams, T.J., Lauro, F.M., Raftery, M., Gibson, J. A.E., et al. (2010) Metaproteogenomic analysis of a dominant green sulfur bacterium from Ace Lake, Antarctica. *ISME J.* **4**: 1002–1019.

Niederberger, T.D., McDonald, I.R., Hacker, A.L., Soo, R.M., Barrett, J.E., Wall, D.H., and Cary, S.C. (2008) Microbial community composition in soils of Northern Victoria Land,

Antarctica. *Environ. Microbiol.* **10**: 1713–1724.

Niederberger, T.D., Sohm, J.A., Gunderson, T.E., Parker, A.E., Tirindelli, J.J., Capone, D.G., et al. (2015) Microbial community composition of transiently wetted Antarctic Dry Valley soils. *Front. Microbiol.* **6**: 9.

Ohyama, Y. and Hiruta, S.I. (1995) The terrestrial arthropods of Sor Røndane in eastern Dronning Maud Land, Antarctica, with biogeographical notes. *Polar Biol.* **15**: 341–347.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., et al. (2015) vegan: Community Ecology Package. R package version 2.3-0. *http://CRAN.R-project.org/package=vegan*.

van Ommen, T.D. and Morgan, V. (2010) Snowfall increase in coastal East Antarctica linked with southwest Western Australian drought. *Nat. Geosci.* **3**: 267–272.

Osanai, Y., Nogi, Y., Baba, S., Nakano, N., Adachi, T., Hokada, T., et al. (2013) Geologic evolution of the Sør Rondane Mountains, East Antarctica: Collision tectonics proposed based on metamorphic processes and magnetic anomalies. *Precambrian Res.* **234**: 8–29.

Osborn, T.J. and Jones, P.D. (2014) The CRUTEM4 land-surface air temperature data set: Construction, previous versions and dissemination via Google earth. *Earth Syst. Sci. Data* **6**: 61–68.

Oswald, G.K.A. and Robin, G. de Q. (1973) Lakes Beneath the Antarctic Ice Sheet. *Nature* **245**: 251–254.

Paolo, F.S., Fricker, H.A., and Padman, L. (2015) Volume loss from Antarctic ice shelves is accelerating. *Science (80-.).* **348**: 327–332.

Parameswaran, P., Jalili, R., Tao, L., Shokralla, S., Gharizadeh, B., Ronaghi, M., and Fire, A.Z. (2007) A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Res.* **35**: 1–9.

Parnikoza, I. (2011) Vascular Plants of the Maritime Antarctic: Origin and Adaptation. *Am. J. Plant Sci.* **02**: 381–395.

Pattyn, F., Matsuoka, K., and Berte, J. (2010) Glacio-meteorological conditions in the vicinity of the Belgian Princess Elisabeth Station, Antarctica. *Antarct. Sci.* **22**: 79.

Pearce, D. A. (2005) The structure and stability of the bacterioplankton community in Antarctic freshwater lakes, subject to extremely rapid environmental change. *FEMS Microbiol. Ecol.* **53**: 61–72.

Pearce, D. A., Bridge, P.D., Hughes, K. A., Sattler, B., Psenner, R., and Russell, N.J. (2009) Microorganisms in the atmosphere over Antarctica. *FEMS Microbiol. Ecol.* **69**: 143–157.

Pearce, D. A., Hughes, K. A., Lachlan-Cope, T., Harangozo, S. A., and Jones, A.E. (2010) Biodiversity of air-borne microorganisms at Halley Station, Antarctica. *Extremophiles* **14**: 145–159.

Pearce, D. A.., Newsham, K.K., Thorne, M. A.S., Calvo-Bado, L., Krsek, M., Laskaris, P., et al. (2012) Metagenomic analysis of a southern maritime Antarctic soil. *Front. Microbiol.* **3**: 1–13.

Pearce, D.A. (2009) Antarctic subglacial lake exploration: a new frontier in microbial ecology. *ISME J.* **3**: 877–880.

Pearce, D.A., Alekhina, I.A., Terauds, A., Wilmotte, A., Quesada, A., Edwards, A., et al. (2016) Aerobiology Over Antarctica – A New Initiative for Atmospheric Ecology. *Front. Microbiol.* **7**: 1–7.

Peat, H.J., Clarke, A., and Convey, P. (2007) Diversity and biogeography of the Antarctic flora. *J. Biogeogr.* **34**: 132–146.

Peck, L.S., Clark, M.S., Clarke, A., Cockell, C.S., Convey, P., Detrich, H.W., et al. (2005) Genomics: Applications to Antarctic ecosystems. *Polar Biol.* **28**: 351–365.

Pedrós-Alió, C. (2007) Ecology. Dipping into the rare biosphere. *Science* **315**: 192–193.

Pedrós-Alió, C. (2012) The Rare Bacterial Biosphere. *Ann. Rev. Mar. Sci.* **4**: 449–466.

Peeters, K., Ertz, D., and Willems, A. (2011) Culturable bacterial diversity at the Princess Elisabeth Station (Utsteinen, Sør Rondane Mountains, East Antarctica) harbours many new taxa. *Syst. Appl. Microbiol.* **34**: 360–367.

Peeters, K., Hodgson, D. A., Convey, P., and Willems, A. (2011) Culturable Diversity of Heterotrophic Bacteria in Forlidas Pond (Pensacola Mountains) and Lundström Lake (Shackleton Range), Antarctica. *Microb. Ecol.* **62**: 399–413.

Peeters, K., Verleyen, E., Hodgson, D. a., Convey, P., Ertz, D., Vyverman, W., and Willems, A. (2011) Heterotrophic bacterial diversity in aquatic microbial mat communities from Antarctica. *Polar Biol.* **35**: 543–554.

Peeters, K. and Willems, A. (2011) The gyrB gene is a useful phylogenetic marker for exploring the diversity of *Flavobacterium* strains isolated from terrestrial and aquatic habitats in Antarctica. *FEMS Microbiol. Lett.* **321**: 130–140.

Peres-Neto, P.R., Legendre, P., Dray, S., and Borcard, D. (2006) Variation partitioning of species data matrices: Estimation and comparison of fractions. *Ecology* **87**: 2614–2625.

Pham, V.H.T. and Kim, J. (2012) Cultivation of unculturable soil bacteria. *Trends Biotechnol.* **30**: 475–484.

Pinto, A.J. and Raskin, L. (2012) PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. *PLoS One* **7**: e43093.

Pointing, S.B. and Belnap, J. (2012) Microbial colonization and controls in dryland systems. *Nat. Rev. Microbiol.* **10**: 551–562.

Pointing, S.B., Chan, Y., Lacap, D.C., Lau, M.C.Y., Jurgens, J. A., and Farrell, R.L. (2009a) Highly specialized microbial diversity in hyper-arid polar desert. *Proc. Natl. Acad. Sci. U. S. A.* **106**: 19964–19969.

Pointing, S.B., Chan, Y., Lacap, D.C., Lau, M.C.Y., Jurgens, J. A., and Farrell, R.L. (2009b) Highly specialized microbial diversity in hyper-arid polar desert. *Proc. Natl. Acad. Sci. U. S. A.* **106**: 19964–19969.

Pollard, D. and DeConto, R.M. (2009) Modelling West Antarctic ice sheet growth and collapse through the past five million years. *Nature* **458**: 329–332.

Popa, R., Popa, R., Mashall, M.J., Nguyen, H., Tebo, B.M., and Brauer, S. (2009) Limitations and benefits of ARISA intra-genomic diversity fingerprinting. *J. Microbiol. Methods* **78**: 111–118.

Powell, J.R. and Sikes, B.A. (2014) Method or madness: Does OTU delineation bias our perceptions of fungal ecology? *New Phytol.* **202**: 1095–1097.

Prosser, J.I. (2015) Dispersing misconceptions and identifying opportunities for the use of "omics" in soil microbial ecology. *Nat. Rev. Microbiol.* **13**: 439–446.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., and Glöckner, F.O. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* **35**: 7188–7196.

Pugh, P.J. A.and Convey, P. (2008) Surviving out in the cold: Antarctic endemic invertebrates and their refugia. *J. Biogeogr.* **35**: 2176–2186.

Pujalte, M.J., Lucena, T., Ruvira, M.A., Arahal, D.R., and Macián, M.C. (2014) The family *Rhodobacteraceae*. In, *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*., pp. 439–512.

Quaiser, A., Ochsenreiter, T., Lanz, C., Schuster, S.C., Treusch, A.H., Eck, J., and Schleper, C. (2003) *Acidobacteria* form a coherent but highly diverse group within the bacterial domain: Evidence from environmental genomics. *Mol. Microbiol.* **50**: 563–575.

Quayle, W.C., Peck, L.S., Peat, H., Ellis-Evans, J.C., and Harrigan, P.R. (2002a) Extreme responses to climate change in Antarctic lakes. *Science* **295**: 645.

Quince, C., Lanzén, A., Curtis, T.P., Davenport, R.J., Hall, N., Head, I.M., et al. (2009) Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat. Methods* **6**: 639–641.

Quince, C., Lanzen, A., Davenport, R.J., and Turnbaugh, P.J. (2011) Removing noise from pyrosequenced amplicons. *BMC Bioinformatics* **12**: 38.

R Core Team (2015) R: A language and environment for statistical computing.

Raymond, J., Siefert, J.L., Staples, C.R., and Blankenship, R.E. (2004) The Natural History of Nitrogen Fixation. *Mol. Biol. Evol.* **21**: 541–554.

Reddy, G.S.N., Aggarwal, R.K., Matsumoto, G.I., and Shivaji, S. (2000) Arthrobacter flavus sp. nov., a psychrophilic bacterium isolated from a pond in McMurdo Dry Valley, Antarctica. *Int. J. Syst. Evol. Microbiol.* **50**: 1553–1561.

Reddy, G.S.N., Matsumoto, G.I., Schumann, P., Stackerbrandt, E., and Shivaji, S. (2004) Psychrophilic pseudomonads from Antarctica: Pseudomonas antartica sp. nov., Pseudomonas meridiana sp. nov. and Pseudomonas proteolytica sp. nov. *Int. J. Syst. Evol. Microbiol.* **54**: 713–719.

Reeder, J. and Knight, R. (2009) The "rare biosphere": a reality check. *Nat. Methods* **6**: 636–7.

Richter, A., Popov, S. V., Schroeder, L., Schwabe, J., Ewert, H., Scheinert, M., et al. (2014) Subglacial Lake Vostok not expected to discharge water. *Geophys. Res. Lett.* **41**: 6772–6778.

Richter, I., Herbold, C.W., Lee, C.K., McDonald, I.R., Barrett, J.E., and Cary, S.C. (2014) Influence of soil properties on archaeal diversity and distribution in the McMurdo Dry Valleys, Antarctica. *FEMS Microbiol. Ecol.* **89**: 347–359.

Robin, G. de Q., Swithinbank, C.W.M., and Smith, B.M.E. (1970) Radio echo exploration of the Antarctic ice sheet. *Int. Assoc. Sci. Hydrol. Publ.* **86**: 97–115.

Ronaghi, M., Karamohamed, S., Pettersson, B., and Uhle, M. (1996) Real-Time DNA Sequencing Using Detection of Pyrophosphate Release. **89**: 84–89.

Ronaghi, M., Uhlén, M., and Nyrén, P. (1998) A sequencing method based on real-time pyrophosphate. *Science* **281**: 363, 365.

Le Roux, P.C. and McGeoch, M.A. (2008) Changes in climate extremes, variability and signature on sub-Antarctic Marion Island. *Clim. Change* **86**: 309–329.

Sabbe, K., Verleyen, E., Hodgson, D.A., Vanhoutte, K., and Vyverman, W. (2003) Benthic diatom flora of freshwater and saline lakes in the Larsemann Hills and Rauer Islands, East Antarctica. *Antarct. Sci.* **15**: 227–248.

Sassenhagen, I., Sefbom, J., Säll, T., Godhe, A., and Rengefors, K. (2015) Freshwater protists do not go with the flow: Population structure in *Gonyostomum semen* independent of connectivity among lakes. *Environ. Microbiol.* **17**: 5063–5072.

Sattley, W.M., Jung, D.O., and Madigan, M.T. (2008) *Psychrosinus fermentans* gen. nov., sp. nov., a lactate-fermenting bacterium from near-freezing oxycline waters of a meromictic Antarctic lake. *FEMS Microbiol. Lett.* **287**: 121–127.

Säwström, C., Granéli, W., Laybourn-Parry, J., and Anesio, A.M. (2007) High viral infection rates in Antarctic and Arctic bacterioplankton. *Environ. Microbiol.* **9**: 250–255.

Schiaffino, M.R., Unrein, F., Gasol, J.M., Massana, R., Balagu??, V., and Izaguirre, I. (2011) Bacterial community structure in a latitudinal gradient of lakes: The roles of spatial versus environmental factors. *Freshw. Biol.* **56**: 1973–1991.

Schloss, P.D. (2010) The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies. *PLoS Comput. Biol.* **6**: e1000844.

Schloss, P.D., Gevers, D., and Westcott, S.L. (2011) Reducing the Effects of PCR Amplification and Sequencing Artifacts on 16S rRNA-Based Studies. *PLoS One* **6**: e27310.

Schloss, P.D. and Handelsman, J. (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl. Environ. Microbiol.* **71**: 1501–1506.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al. (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**: 7537–7541.

Sekiguchi, Y., Yamada, T., Hanada, S., Ohashi, A., Harada, H., and Kamagata, Y. (2003) *Anaerolinea thermophila* gen. nov., sp. nov. and *Caldilinea aerophila* gen. nov., sp. nov.,

novel filamentous thermophiles that represent a previously uncultured lineage of the domain bacteria at the subphylum level. *Int. J. Syst. Evol. Microbiol.* **53**: 1843–1851.

Sellstedt, A. and Richau, K.H. (2013) Aspects of nitrogen-fixing *Actinobacteria*, in particular free-living and symbiotic *Frankia*. *FEMS Microbiol. Lett.* **342**: 179–186.

Seneviratne, G. and Indrasena, I.K. (2006) Nitrogen fixation in lichens is important for improved rock weathering. *J. Biosci.* **31**: 639–643.

Shade, A., Hogan, C.S., Klimowicz, A.K., Linske, M., McManus, P.S., and Handelsman, J. (2012) Culturing captures members of the soil rare biosphere. *Environ. Microbiol.* **14**: 2247–52.

Sheridan, P.P., Loveland-Curtze, J., Miteva, V.I., and Brenchley, J.E. (2003) *Rhodoglobus vestalii* gen. nov., sp. nov., a novel psychrophilic organism isolated from an Antarctic Dry Valley Lake. *Int. J. Syst. Evol. Microbiol.* **53**: 985–994.

Shivaji, S., Kumari, K., Kishore, K.H., Pindi, P.K., Rao, P.S., Radha Srinivas, T.N., et al. (2011) Vertical distribution of bacteria in a lake sediment from Antarctica by culture-independent and culture-dependent approaches. *Res. Microbiol.* **162**: 191–203.

Shivaji, S., Reddy, G.S.N., Suresh, K., Gupta, P., Chintalapati, S., Schumann, P., et al. (2005) *Psychrobacter vallis* sp. nov. and *Psychrobacter aquaticus* sp. nov., from Antarctica. *Int. J. Syst. Evol. Microbiol.* **55**: 757–762.

Siegert, M.J. (2000) Antarctic subglacial lakes. *Earth Sci. Rev.* **50**: 29–50.

Siegert, M.J., Carter, S., Tabacco, I., Popov, S., and Blankenship, D.D. (2005) A revised inventory of Antarctic subglacial lakes. *Antarct. Sci.* **17**: 453.

Siegert, M.J., Dowdeswell, J. A., Gorman, M.R., and McIntyre, N.F. (1996) An inventory of Antarctic sub-glacial lakes. *Antarct. Sci.* **8**: 281–286.

Sipos, R., Székely, A.J., Palatinszky, M., Révész, S., Márialigeti, K., and Nikolausz, M. (2007) Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis. *FEMS Microbiol. Ecol.* **60**: 341–350.

Smith, D.J., Timonen, H.J., Jaffe, D. A., Griffin, D.W., Birmele, M.N., Perry, K.D., et al. (2013) Intercontinental dispersal of bacteria and archaea by transpacific winds. *Appl. Environ. Microbiol.* **79**: 1134–9.

Smith, J.J., Tow, L.A., Stafford, W., Cary, C., and Cowan, D. A. (2006) Bacterial diversity in three different antarctic cold desert mineral soils. *Microb. Ecol.* **51**: 413–421.

Smith, J.L., Barrett, J.E., Tusnády, G., Rejtö, L., and Cary, S.C. (2010) Resolving environmental drivers of microbial community structure in Antarctic soils. *Antarct. Sci.* **22**: 673–680.

Smith, R.I.L. (1994) Vascular plants as bioindicators of regional warming in Antarctica. *Oecologia* **99**: 322–328.

Sogin, M.L., Morrison, H., Mclellan, S., Welch, D.M., and Huse, S. (2010) The rare biosphere : sorting out fact from fiction. **11**: 2010.

Sogin, M.L., Sogin, M.L., Morrison, H.G., Morrison, H.G., Huber, J. A., et al. (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc. Natl. Acad. Sci. U. S. A.* **103**: 12115–20.

Sokol, E.R., Herbold, C.W., Lee, C.K., Cary, S.C., and Barrett, J.E. (2013) Local and regional influences over soil microbial metacommunities in the Transantarctic Mountains. *Ecosphere* **4**: 1–24.

Soo, R.M., Wood, S. A., Grzymski, J.J., McDonald, I.R., and Cary, S.C. (2009) Microbial biodiversity of thermophilic communities in hot mineral soils of Tramway Ridge, Mount Erebus, Antarctica. *Environ. Microbiol.* **11**: 715–728.

Staley, J.T. and Konopka, A. (1985) Measurement of in Situ Activities of Nonphotosynthetic Microorganisms in Aquatic and Terrestrial Habitats. *Annu. Rev. Microbiol.* **39**: 321–346.

Steig, E.J. and Orsi, A.J. (2013) The heat is on in Antarctica. *Nat. Geosci.* **6**: 6–8.

Stevens, M.I. and D'Haese, C.A. (2014) Islands in ice: isolated populations of *Cryptopygus sverdrup* (Collembola) among nunataks in the Sør Rondane Mountains, Dronning Maud Land, Antarctica. *Biodiversity* **15**: 169–177.

Stevens, M.I., Greenslade, P., Hogg, I.D., and Sunnucks, P. (2006) Southern hemisphere springtails: Could any have survived glaciation of antarctica? *Mol. Biol. Evol.* **23**: 874–882.

Stewart, E.J. (2012) Growing Unculturable Bacteria. *J. Bacteriol.* **194**: 4151–4160.

Stomeo, F., Makhalanyane, T.P., Valverde, A., Pointing, S.B., Stevens, M.I., Cary, C.S., et al. (2012) Abiotic factors influence microbial diversity in permanently cold soil horizons of a maritime-associated Antarctic Dry Valley. *FEMS Microbiol. Ecol.* **82**: 326–340.

Straka, R.P. and Stokes, J.L. (1960) Psychrophilic bacteria from Antarctica. *J. Bacteriol.* **80**: 622–625.

Strunecký, O., Elster, J., and Komárek, J. (2012) Molecular clock evidence for survival of Antarctic cyanobacteria (*Oscillatoriales, Phormidium autumnale*) from Paleozoic times. *FEMS Microbiol. Ecol.* **82**: 482–90.

Sumner, D.Y., Hawes, I., Mackey, T.J., Jungblut, A.D., and Doran, P.T. (2015) Antarctic microbial mats: A modern analog for Archean lacustrine oxygen oases. *Geology* **43**: 887–890.

Swenson, U., Hill, R., and McLoughlin, S. (2001) Biogeography of *Nothofagus* supports the sequence of Gondwana break-up. *Taxon* **50**: 1025–1041.

Tahon, G., Tytgat, B., Stragier, P., and Willems, A. (2016) Analysis of cbbL, nifH, and pufLM in Soils from the Sør Rondane Mountains, Antarctica, Reveals a Large Diversity of Autotrophic and Phototrophic Bacteria. *Microb. Ecol.* **71**: 131–149.

Tanaka, T., Kawasaki, K., Daimon, S., Kitagawa, W., Yamamoto, K., Tamaki, H., et al. (2014) A Hidden Pitfall in the Preparation of Agar Media Undermines Microorganism Cultivability. *Appl. Environ. Microbiol.* **80**: 7659–7666.

Taton, A., Grubisic, S., Brambilla, E., De Wit, R., and Wilmotte, A. (2003) Cyanobacterial

Diversity in Natural and Artificial Microbial Mats of Lake Fryxell (McMurdo Dry Valleys, Antarctica): a Morphological and Molecular Approach. *Appl. Environ. Microbiol.* **69**: 5157–5169.

Tavernier, I., Verleyen, E., Hodgson, D. a, Heirman, K., Roberts, S.J., Imura, S., et al. (2014) Absence of a Medieval Climate Anomaly, Little Ice Age and twentieth century warming in Skarvsnes, Lutzow Holm Bay, East Antarctica. *Antarct. Sci.* **26**: 585–598.

Teixeira, L.C.R.S., Peixoto, R.S., Cury, J.C., Sul, W.J., Pellizari, V.H., Tiedje, J., and Rosado, A.S. (2010) Bacterial diversity in rhizosphere soil from Antarctic vascular plants of Admiralty Bay, maritime Antarctica. *ISME J.* **4**: 989–1001.

Temperton, B. and Giovannoni, S.J. (2012) Metagenomics: microbial diversity through a scratched lens. *Curr. Opin. Microbiol.* **15**: 605–612.

Terauds, A., Chown, S.L., Morgan, F., Peat, H.J., Watts, D.J., Keys, H., et al. (2012) Conservation biogeography of the Antarctic. *Divers. Distrib.* **18**: 726–741.

The cultural revolution (2012) *Nat. Rev. Microbiol.* **11**: 1–1.

Thompson, D.W.J. and Solomon, S. (2002) Interpretation of recent Southern Hemisphere climate change. *Science (80-.).* **296**: 895–899.

Tindall, B.J. (2004) Prokaryotic diversity in the Antarctic: the tip of the iceberg. *Microb. Ecol.* **47**: 271–283.

Van Trappen, S., Mergaert, J., Van Eygen, S., Dawyndt, P., Cnockaert, M.C., and Swings, J. (2002) Diversity of 746 heterotrophic bacteria isolated from microbial mats from ten Antarctic lakes. *Syst. Appl. Microbiol.* **25**: 603–610.

Van Trappen, S., Vandecandelaere, I., Mergaert, J., and Swings, J. (2004) *Gillisia limnaea* gen. nov., sp. nov., a new member of the family Flavobacteriaceae isolated from a microbial mat in Lake Fryxell, Antartica. *Int. J. Syst. Evol. Microbiol.* **54**: 445–448.

Trivelpiece, W.Z., Hinke, J.T., Miller, A.K., Reiss, C.S., Trivelpiece, S.G., and Watters, G.M. (2011) Variability in krill biomass links harvesting and climate warming to penguin population changes in Antarctica. *Proc. Natl. Acad. Sci.* **108**: 7625–7628.

Tsujimoto, M., McInnes, S.J., Convey, P., and Imura, S. (2014) Preliminary description of tardigrade species diversity and distribution pattern around coastal Syowa Station and inland Sør Rondane Mountains, Dronning Maud Land, East Antarctica. *Polar Biol.* **37**: 1361–1367.

Turner, J., Bindschadler, R., Convey, P., Prisco, G. Di, Fahrbach, E., Gutt, J., et al. (2009) Antarctic Climate Change and the Environment.

Tytgat, B., Verleyen, E., Obbels, D., Peeters, K., De Wever, A., D'hondt, S., et al. (2014) Bacterial diversity assessment in antarctic terrestrial and aquatic microbial mats: A comparison between bidirectional pyrosequencing and cultivation. *PLoS One* **9.**:

Ugolini, F.C. and Bockheim, J.G. (2008) Antarctic soils and soil formation in a changing environment: A review. *Geoderma* **144**: 1–8.

Valverde, A., Makhalanyane, T.P., Seely, M., and Cowan, D. A. (2015) Cyanobacteria drive

community composition and functionality in rock-soil interface communities. *Mol. Ecol.*

Varin, T., Lovejoy, C., Jungblut, A.D., and Vincent, W.F. (2010) Metagenomic profiling of Arctic microbial mat communities as nutrient scavenging and recycling systems. **55**: 1901–1911.

Varin, T., Lovejoy, C., Jungblut, A.D., Vincent, W.F., and Corbeil, J. (2012) Metagenomic analysis of stress genes in microbial mat communities from Antarctica and the high Arctic. *Appl. Environ. Microbiol.* **78**: 549–559.

Veillette, J., Lovejoy, C., Potvin, M., Harding, T., Jungblut, A.D., Antoniades, D., et al. (2011) Milne Fiord Epishelf Lake: A Coastal Arctic Ecosystem Vulnerable to Climate Change. *Ecoscience* **18**: 304–316.

Veillette, J., Mueller, D.R., Antoniades, D., and Vincent, W.F. (2008) Arctic epishelf lakes as sentinel ecosystems: Past, present and future. *J. Geophys. Res.* **113**: G04014.

Velasco-Castrillón, A., Gibson, J. A. E., and Stevens, M.I. (2014) A review of current Antarctic limno-terrestrial microfauna. *Polar Biol.*

Verleyen, E., Hodgson, D.A., Vyverman, W., Roberts, D., McMinn, A., Vanhoutte, K., and Sabbe, K. (2003) Modelling diatom responses to climate induced fluctuations in the moisture balance in continental Antarctic lakes. In, *Journal of Paleolimnology*., pp. 195–215.

Verleyen, E., Sabbe, K., Hodgson, D. A., Grubisic, S., Taton, A., Cousin, S., et al. (2010) Structuring effects of climate-related environmental factors on Antarctic microbial mat communities. *Aquat. Microb. Ecol.* **59**: 11–24.

Verleyen, E., Vyverman, W., Sterken, M., Hodgson, D. A., De Wever, A., Juggins, S., et al. (2009) The importance of dispersal related and local factors in shaping the taxonomic structure of diatom metacommunities. *Oikos* **118**: 1239–1249.

Vincent, W.F. (2000) Evolutionary origins of Antarctic microbiota: invasion, selection and endemism. *Antarct. Sci.* **12**: 374–385.

Vincent, W.F., Gibson, J. A.E., Pienitz, R., Villeneuve, V., Broady, P. A. Hamilton, P.B., and Howard-Williams, C. (2000) Ice shelf microbial ecosystems in the high arctic and implications for life on snowball earth. *Naturwissenschaften* **87**: 137–141.

Vrionis, H.A., Whyte, L.G., and Miller, R. V. (2013) Life at the Poles in the Age of Global Warming: Part 2. *Microbe* **8**: 491.

Vyverman, W., Verleyen, E., Sabbe, K., Vanhoutte, K., Sterken, M., Hodgson, D. A., et al. (2007) Historical processes constrain patterns in global diatom diversity. *Ecology* **88**: 1924–31.

Vyverman, W., Verleyen, E., Wilmotte, A., Hodgson, D. A., Willems, A., Peeters, K., et al. (2010) Evidence for widespread endemism among Antarctic micro-organisms. *Polar Sci.* **4**: 103–113.

Wada, E., Shibata, R., and Torii, T. (1981) 15N abundance in Antarctica: origin of soil nitrogen and ecological implications. *Nature* **292**: 327–329.

Walker, J.J. and Pace, N.R. (2007) Endolithic Microbial Ecosystems. *Annu. Rev. Microbiol.* **61**: 331–347.

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., et al. (2002) Ecological responses to recent climate change. *Nature* **426**: 389–395.

Wang, N.F., Zhang, T., Zhang, F., Wang, E.T., He, J.F., Ding, H., et al. (2015) Diversity and structure of soil bacterial communities in the Fildes Region (maritime Antarctica) as revealed by 454 pyrosequencing. *Front. Microbiol.* **6.**:

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**: 5261–7.

Ward, N., Challacombe, J., Janssen, P., and Henrissat, B. (2009) Three genomes from the phylum *Acidobacteria* provide insight into their lifestyles in soils. *Appl. Environ. Microbiol.*

Waters, J.M. and Craw, D. (2006) Goodbye Gondwana? New Zealand biogeography, geology, and the problem of circularity. *Syst. Biol.* **55**: 351–356.

Werner, J.J., Koren, O., Hugenholtz, P., DeSantis, T.Z., Walters, W. A., Caporaso, J.G., et al. (2012) Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. *ISME J.* **6**: 94–103.

Whitaker, D. and Christman, M. (2014) clustsig: Significant Cluster Analysis.

Whitaker, R.J., Grogan, D.W., and Taylor, J.W. (2003) Hyperthermophilic *Archaea*. *Science (80-.).* **301**: 2002–2004.

Wilkins, D., Yau, S., Williams, T.J., Allen, M.A., Brown, M. V, Demaere, M.Z., et al. (2013) Key microbial drivers in Antarctic aquatic environments. *FEMS Microbiol. Rev.* **37**: 303–335.

Willig, M., Kaufman, D., and Stevens, R. (2003) Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annu. Rev. Ecol. Evol. Syst.* **34**: 273–309.

Wingham, D.J., Siegert, M.J., Shepherd, A., and Muir, A.S. (2006) Rapid discharge connects Antarctic subglacial lakes. *Nature* **440**: 1033–1036.

Wood, S. A., Mountfort, D., Selwood, A.I., Holland, P.T., Puddick, J., and Cary, S.C. (2008) Widespread distribution and identification of eight novel microcystins in antarctic cyanobacterial mats. *Appl. Environ. Microbiol.* **74**: 7243–51.

Wood, S. A., Rueckert, A., Cowan, D. A., and Cary, S.C. (2008) Sources of edaphic cyanobacterial diversity in the Dry Valleys of Eastern Antarctica. *ISME J.* **2**: 308–20.

Wright, A. and Siegert, M. (2012) A fourth inventory of Antarctic subglacial lakes. *Antarct. Sci.* **24**: 659–664.

Wynn-Williams, D.D. (1996) Antarctic microbial diversity: the basis of polar ecosystem processes. *Biodivers. Conserv.* **5**: 1271–1293.

Yarwood, S., Wick, A., Williams, M., and Daniels, W.L. (2014) Parent Material and Vegetation Influence Soil Microbial Community Structure Following 30-Years of Rock

Weathering and Pedogenesis. *Microb. Ecol.* **69**: 383–394.

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.-H., et al. (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol.* **12**: 635–645.

Yau, S., Lauro, F.M., Williams, T.J., Demaere, M.Z., Brown, M. V, Rich, J., et al. (2013) Metagenomic insights into strategies of carbon conservation and unusual sulfur biogeochemistry in a hypersaline Antarctic lake. *ISME J.* **7**: 1944–61.

Yergeau, E., Bokhorst, S., Kang, S., Zhou, J., Greer, C.W., Aerts, R., and Kowalchuk, G. A. (2012) Shifts in soil microorganisms in response to warming are consistent across a range of Antarctic environments. *ISME J.* **6**: 692–702.

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

Yergeau, E., Newsham, K.K., Pearce, D.A., and Kowalchuk, G.A. (2007) Patterns of bacterial diversity across a range of Antarctic terrestrial habitats. *Environ. Microbiol.* **9**: 2670–82.

Yilmaz, P., Kottmann, R., Field, D., Knight, R., Cole, J.R., Amaral-zettler, L., et al. (2011) perspective Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. **29**: 415–420.

Youssef, N., Sheik, C.S., Krumholz, L.R., Najar, F.Z., Roe, B. A., and Elshahed, M.S. (2009) Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. *Appl. Environ. Microbiol.* **75**: 5227–5236.

Youssef, N., Steidley, B.L., and Elshahed, M.S. (2012) Novel High-Rank Phylogenetic Lineages within a Sulfur Spring (Zodletone Spring, Oklahoma), Revealed Using a Combined Pyrosequencing-Sanger Approach. *Appl. Environ. Microbiol.* **78**: 2677–2688.

Yu, Z. and Morrison, M. (2004) Comparisons of different hypervariable regions of rrs genes for use in fingerprinting of microbial communities by PCR-denaturing gradient gel electrophoresis. *Appl. Environ. Microbiol.* **70**: 4800–4806.

Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014) PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**: 614–620.

Zhang, L., Jungblut, A.D., Hawes, I., Andersen, D.T., Sumner, D.Y., and Mackey, T.J. (2015) Cyanobacterial diversity in benthic mats of the McMurdo Dry Valley lakes, Antarctica. *Polar Biol.* **38**: 1097–1110.

Zhou, J., Jiang, Y.H., Deng, Y., Shi, Z., Zhou, B.Y., Xue, K., et al. (2013) Random sampling process leads to overestimation of β-diversity of microbial communities. *MBio* **4**: e00324–13–e00324–13.

Zwart, G., Hiorns, W.D., Methé, B. A., van Agterveld, M.P., Huismans, R., Nold, S.C., et al. (1998) Nearly identical 16S rRNA sequences recovered from lakes in North America and Europe indicate the existence of clades of globally distributed freshwater bacteria. *Syst. Appl. Microbiol.* **21**: 546–556.

Curriculum vitae

Personalia

Bjorn Tytgat
Opaaigem 28, 9420 Erpe-Mere, Belgium
Born: February 11th, 1980, Izegem, Belgium
Bjorn.Tytgat@UGent.be

Married to Wendy Van den Steen (Msc. Vet. Sci.), one daughter, Linnéa (°2014)

Professional experience

March 17th 2011 – 30th June 2012: Project: *State-of-the-art techniques for the study of microbial diversity patterns in East Antarctica*. LM-UGent, Ghent University

July 1st 2012 – 30th June 2016: Project *CCAMBIO: Climate Change and Antarctic Microbial Biodiversity*. LM-UGent, Ghent University

Educational background

2011-present	PhD-student, LM-UGent, Department of Microbiology and Biochemistry, Faculty of Science, Ghent University Dissertation: Distribution and characterization of bacterial communities in diverse Antarctic ecosystems by high-throughput sequencing. Promotors: Prof. Dr. Anne Willems, Prof. Dr. Elie Verleyen
2007-2009	Ghent University, Ghent, Belgium Degree of Master of Science in Biology Dissertation: Evolution of floating in tropical passerines: a theoretical approach. Promotors: Prof. Dr. Dries Bonte, Prof. Dr. Luc Lens
2004-2007	Ghent University, Ghent, Belgium Degree of Bachelor of Science in Biology Dissertation: Effects of habitat fragmentation on the dispersal capacity of some wing-polymorphic carabids in European agricultural areas. Promotors: Dr. Frederik Hendrickx, Dr. Konjev Desender
2003-2004	Ghent University, Ghent, Belgium 1st year candidature Bioengineering
2002-2003	Examination Committee for Secondary Education (Tweedekansonderwijs), Mathematics-Sciences, Kortrijk & Bruges
1999-2001	Ter Groene Poorte, Bruges Small Business Management & Entrepreneurship: Bakery

1996-1999 Centrum voor Middenstandsopleiding, Kortrijk
Apprenticeship baker (leercontract bakkerij)

1992-1996 Rijksmiddenschool Bellevue, Izegem
Secondary education: Latin, Economics-Modern Languages

Language proficiencies

Dutch – native speaker
English – excellent command
French – good command
German – good working knowledge
Swedish – working knowledge
Japanese – basic communication skills; reading: katakana & hiragana: good, kanji: basic
Breton, Turkish, Spanish, Italian, Norwegian: notions

Programming languages and computer skills

Proficient: R, Linux, Microsoft Windows and Office, html & css, Google
Familiar with: Python, Delphi, SQL

Scientific output

A1 publications

1. Tytgat, B., Verleyen, E., Obbels, D., Peeters, K., De Wever, A., D'hondt, S., De Meyer, T., Van Criekinge, W., Vyverman, W. and Willems, A. (2014) Bacterial diversity assessment in Antarctic terrestrial and aquatic microbial mats: a comparison between bidirectional pyrosequencing and cultivation. PLoS One. 9. doi:10.1371/journal.pone.0097564

2. Decleyre, H., Heylen, K., Sabbe, K., Tytgat, B., Deforce, D., Van Nieuwerburgh, F., Van Colen, C. and Willems, A. (2015). A Doubling of Microphytobenthos Biomass Coincides with a Tenfold Increase in Denitrifier and Total Bacterial Abundances in Intertidal Sediments of a Temperate Estuary. PLoS One.10: e0126583. doi:10.1371/journal.pone.0126583

3. Tahon, G., Tytgat, B., Stragier, P. and Willems A. (2016). Analysis of RuBisCO, nifH and pufLM genes in soils near the Princess Elisabeth Station, Sør Rondane Mountains, Antarctica, indicates a large diversity of auto- and phototrophic bacteria. Microb Ecol. doi:10.1007/s00248-015-0704-6

4. Decleyre, H., Heylen, K., Tytgat, B. and Willems A. (accepted). Highly diverse *nirK* genes comprise two major clades that harbor ammonium-producing denitrifiers. BMC Genomics.

5. Obbels, D., Verleyen, E., Mano, M-J., Namsaraev, Z., Sweetlove, M., Tytgat, B., Fernandez-Carazo, R., De Wever, A., D'hondt, S., Ertz, D., Elster, J., Sabbe, K., Willems, A., Wilmotte, A., and Vyverman, W. Prokaryotic and eukaryotic biodiversity patterns in

terrestrial habitats of the Sør Rondane Mountains, Dronning Maud Land, East Antarctica. FEMS Microbiology Ecology- *accepted*.

6. Tytgat, B., Verleyen, E., Sweetlove, M., D'hondt, S., Clercx, P., Van Ranst, E., Peeters, K., Roberts, S., Namsaraev, Z., Wilmotte, A., Vyverman, W. and Willems A. Bacterial community composition in relation to bedrock type and macrobiota in soils from the Sør Rondane Mountains, East Antarctica. *Submitted* to FEMS Microbiology Ecology.

7. Tahon, G., Tytgat, B. and Willems A. Illumina paired-end sequencing reveals diversity of key genes for primary production and diazotrophy in soils from the Sør Rondane Mountains, East-Antarctica. *Submitted BMC Genomics.*

A4 publications

1. Verleyen E, Obbels D, De Wever A, Souffreau C, Vanormelingen P, Sabbe K, Vyverman W, Peeters K, Tytgat B, Willems A, Ertz D, Van de Vijver B, Mano M-J, De Carvalho Maalouf P, Fernández-Carazo R, Namsaraev Z & Wilmotte A (2013) Out of sight, out of mind? Antarctic microbial diversity as an additional criterion for conservation purposes. Science Connection 41: 44-47

2. Tytgat B (2014) Zeg niet zomaar een braam. Dendriet, tijdschrift van Natuurpunt Denderstreek. 1-2014: 26-30. http://issuu.com/dendriet/docs/dendriet_jrg13_nr1_2014

Conferences (talks and posters)

1. Tytgat B & Willems A. *Comparing Antarctic microbial communities: cultivation vs. pyrosequencing*. APECS meeting - Belgian Young Researchers Antarctica Day workshop, 1 December, 2011, Brussels, Belgium
2. Tytgat B, Verleyen E, Obbels D, D'hondt S, Peeters K, Vyverman W and Willems A. *Microbial Diversity patterns in Antarctica – processing pyrosequencing data*. FWO Discussion Meeting on Characterizing bacterial communities: comparing first- and next-generation molecular techniques. 26th April 2012, Leuven, Belgium
3. Decleyre H, Heylen K, Tytgat B, Van Colen C and Willems A (2012) *Diversity and abundance of dissimilatory nitrogen reducers present in the sediments of the Paulina polder tidal mudflat, the Netherlands*. 17th European Nitrogen-Cycle Meeting, Oslo, Norway
4. Vyverman W, Verleyen E, Obbels D, Tytgat B, Wilmotte A, Willems A, Van Nieuwenhuyze W, Tavernier I, Hodgson DA & Sabbe K. *The imprint of glacial history on the biogeography of Antarctic lake-dwelling micro-organisms.* 5th International Conference on Polar and Alpine Microbiology, 08-12 September 2012, Big Sky, Montana, USA.
5. Convey P, Danis B, Laughinghouse IV HD, Obbels D, Pearce D, Pessi IS, Tytgat B, Van de Vijver B, Verleyen E, Vyverman W, Willems A & Wilmotte A. *The CCAMBIO project: responses of the aquatic microbial mats to Climate Change*. Next Generation Sequencing at the Poles, 21-23 November 2012, Liège, Belgium.

6. Decleyre H, Heylen K, Tytgat B, Van Colen C and Willems A (2013) *Diversity and abundance of dissimilatory nitrate reducers present in estuarine sediments of the Paulina Polder tidal flat, The Netherlands*. 18th European Nitrogen Cycle Meeting, Darmstadt, Germany
7. Obbels D, Verleyen E, Tytgat B, Elster J, Strunecky O, Wilmotte A, Willems A, Sabbe K & Vyverman W. *The diversity and tolerance to osmotic stress of East Antarctic filamentous* Cyanobacteria. XI SCAR Biology Symposium**,** 15-19 July 2013, Barcelona, Spain.
8. Tytgat B, Verleyen E, Obbels D, Peeters K, De Wever A, D'hondt S, De Meyer T, Van Criekinge W, Vyverman W & Willems A *Bacterial Diversity Assessment in Antarctic Microbial Mats - A comparison between bidirectional pyrosequencing and cultivation*, Metagenomics Symposium, 19th September 2014, Wageningen, The Netherlands
9. Tahon G, Tytgat B, Stragier P and Willems A. *Life on the frozen continent: diversity of Rubisco, nifH and pufLM genes in soils around the Princess Elisabeth Station, Sør Rondane Mountains, Antarctica.* BSM meeting, 18 November 2014, Brussels, Belgium (poster)
10. Heylen K, Decleyre H, Sabbe K, Tytgat B, Van Colen C and Willems A (2015) *Microphytobenthos increases denitrifier and total bacterial abundances in intertidal sediments of a temperate estuary*. VLIZ Young Marine Scientists' Day 2015, Oostende, Belgium
11. Wilmotte A, Stelmach-Pessi I, Sweetlove M, Obbels D, Vanormelingen P, Tytgat B, Willems A, Verleyen E, Vyverman W, Van De Vijver B, Lara Y and Van de Putte A. *Microorganisms in Antarctic Lacustrine Microbial Mats*. Aquatic Science, Aquatic Sciences Meeting, 25 February 2015, Granada, Spain
12. Tytgat B, Verleyen, E, D'hondt S, Clercx P, Peeters K, Van Ranst E, Vyverman W and Willems A. *Exploring microbial diversity patterns in the Sør Rondane Mountains (East-Antarctica) using Next Generation Sequencing and ARISA*. FEMS 2015, June 2015, Maastricht, the Netherlands. (poster)
13. Tytgat B, Verleyen E, Sweetlove M, Obbels D, D'hondt S, Wilmotte A, Vyverman W and Willems A. *Bacterial community composition in relation to bedrock type and macrobiota in soils from the Sør Rondane Mountains, East Antarctica*. Polar and Alpine Microbiology meeting, 6th International Conference on Polar and Alpine Microbiology, 6-10 September 2015, České Budějovice, Czech Republic. (poster)
14. Tytgat B, Verleyen E, Sweetlove M, Obbels D, D'hondt S, Wilmotte A, Vyverman W and Willems A. *Biogeographic patterns in Antarctic lacustrine prokaryotes*. Polar and Alpine Microbiology meeting, 6th International Conference on Polar and Alpine Microbiology, 6-10 September 2015, České Budějovice, Czech Republic. (poster)
15. Tahon G, Tytgat B and Willems A. *Diversity of cbbL, nifH and pufM genes in soils around the Princess Elisabeth Station, Sør Rondane Mountains, Antarctica*. Polar and Alpine Microbiology meeting, 6th International Conference on Polar and Alpine Microbiology, 6-10 September 2015, České Budějovice, Czech Republic. (presentation)
16. Verleyen E, Vyverman W, Pinseel E, Kopalová K, Antoniades D, Sterken M, Nedbalová L, Jones VJ, Tavernier I, Tytgat B, Souffreau C, Imura S, Kudoh S,

Convey P, Hodgson DA, Sabbe K, Van de Vijver B. *Post-Miocene divergence of polar diatom biomes*. Polar and Alpine Microbiology meeting, 6th International Conference on Polar and Alpine Microbiology, 6-10 September 2015, České Budějovice, Czech Republic. (presentation)

17. Sweetlove M, Verleyen E, Tytgat B, Obbels D, D'hondt S, Willems A. and Vyverman W. *Biogeographic zoning of aquatic microeukaryotes in the Antarctic Region*. Polar and Alpine Microbiology meeting, 6th International Conference on Polar and Alpine Microbiology, 6-10 September 2015, České Budějovice, Czech Republic. (presentation)

Courses followed

1. Getting started with high-performance computing, 30th May-6th June 2012. Ghent University Doctoral Schools training.
2. Metagenomic Methods for Microbial Ecologists, 15-19th September 2014. NIOO, Wageningen, the Netherlands.

Supervision of students and practical courses

1. Yao Adjiguita Kolombia (2011) Master project Biotechnology and Biochemistry
2. Tine Duthoit (2011) Master project Biotechnology and Biochemistry
3. Teaching assistant of the Mothur software suite at the CCAMBIO workshop 'Next-Generation Sequencing at the Poles', 22-23 November 2012, Liège, Belgium
4. Sam Lambrechts (2012-2013) *Exploring the unclassified fraction and the Deinococcales in biofilms from Antarctica by pyrosequencing and genotypic fingerprints* – Master thesis
5. Brecht Bruneel (2013) Master project Biotechnology and Biochemistry
6. Jonas Van Damme (2013) Master project Biotechnology and Biochemistry
7. Hanneloor Heynderickx (2014) *Bacterial community analysis in aquatic microbial mats in Antarctica by pyrosequencing* – Master thesis
8. Practical course Microbiology, 2nd Bachelor Biology (2012 – 2015)
9. Integrated practical course Biotechnology and Biochemistry (2014-2015) partim Bioinformatics: Primer design

Events

1. Co-organizing Dispersal Symposium, 14-15 September 2009, Ghent, Belgium
2. Co-organizing CCAMBIO workshop 'Next-Generation Sequencing at the Poles', November 21-23 2012, Liège, Belgium

Scientific awards

Best poster presentation of early career scientists at the 6th International Conference on Polar and Alpine Microbiology. September 6-10, 2015, České Budějovice, Czech Republic.

Title: Biogeographic patterns in Antarctic lacustrine prokaryotes.