Investigation into the microbial ecology and persistence of soil taxa through trace-gas oxidation in Tongariro National Park

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

Master of Science

at

The University of Waikato

by

HANNA ANNETTE CAROLINE GEORGIA PEACH

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2019

Abstract

Cultivation-dependent and metagenomic surveys of bare, volcanic and hyperarid oligotrophic ecosystems have revealed surprisingly diverse microbial communities[1-4]. The richness of Mt. Tongariro soil communities is surprising considering the extreme environmental parameters. Given that soil moisture on Mt. Tongariro is highly variable dependent on season, and both phototrophic Cyanobacteria and trace gas oxidising taxa from the phyla Actinobacteria, WPS-2 and AD3 are detected in these soils, it is postulated that these taxa alternate as the primary producers in these oligotrophic soils. In this scenario, cyanobacteria are the phototrophic backbone and trace-gas oxidisers use atmospheric trace gases (H_2, CO) as supplemental cell energy for mixotrophic metabolism to persist. Mt. Tongariro soil communities were shown to encode high affinity hydrogenases and carbon monoxide dehydrogenases, which are necessary for atmospheric oxidation of H_2 and CO. This thesis details the microbial ecology and effect of physicochemistry on the alpha and beta diversity of Mt. Tongariro soils. It also addresses microbial persistence by the use oxidation of trace gases H_2 and CO at 4 °C and 60 °C for the South, Red and Central Craters of Mt. Tongariro on the southernmost point of the Taupo Volcanic Centre (TVZ), and control soils from White Island (Whakaari), Mt. Urchin and Mt. Sugarloaf. This preliminary study uncovers the microbial ecology of Mt. Tongariro and other bare, high elevation control soils from around New Zealand. Although, more extensive testing is required such as testing isolated Mt. Tongariro strains to understand whether the thermophilic taxa present are capable of oxidising trace gas, or whether they harbour low affinity enzymes suited to elevated levels of CO and H_2. This study indicates that trace gas oxidation may be an important mechanism for persistence or a dependable source of supplemental energy to specific mixotrophic trace-gas oxidising taxa in all communities from bare, oligotrophic systems.

Acknowledgements

First, I would like to acknowledge my primary supervisor: Professor Ian MacDonald, I thank you for reading my numerous thesis drafts and giving me valuable comments on the structure and content of the document. Most of all, I would like to thank you for being present, understanding and dependable in times of need, especially in helping me gain extensions to cover me due to my hospitalisations over the past year.

Second, I would like to thank both of my secondary supervisors: Dr Matthew Stott and Dr Carlo Carere. Thank you for your continual, nurturing support. I would also like to thank you both for the advice on how to structure these experiments and for taking the time to sit down and bounce around ideas over morning cups of coffee. Also, I would like to thank you both for sifting through my numerous thesis drafts and the sage advice you have given me over the past four years. The encouragement and mentoring you have both given me has allowed me space to cultivate my own style in research with the comfort that I have great support if needed. Additionally, I would like to thank you both, especially Matthew for allowing me to complete this Master of Science with support from GNS Science.

To my current colleagues, Jean, Karen and Lucy: Thank you for the guidance, extensive lab support, friendship and inspiration. It has been truly wonderful to have three strong female influences over the past four years at GNS Science. I would especially like to thank Jean for her extensive help with bioinformatics and for the crash course in statistical analyses for ecological data. I would like to thank Karen for answering all my questions with kindness. To the honorary bugs: I would like to thank Lucjan for his caring friendship, daily chats, gifts of chocolate and baking over the past few years. To an external honorary bug, Dr Kevin Lee, I would like to thank you for your preliminary help with bioinformatics and for the friendship over the past four years.

Lastly, I would like to thank my partner, Robert, and my close friends from outside of work who have encouraged and supported me throughout this entire learning process.

Table of Contents

List of Figures

List of Tables

List of Equations

Abbreviations

S.I (Systeme International d'unites) abbreviations for units and standard notations for chemicals, elements, formulae, and chemical abbreviations are used in this thesis. All additional acronyms are detailed below.

ASV	Assigned Sequence Variance
BLAST	Basic Local Alignment Tool
bp	Base pair
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphate
dsDNA	Double-stranded deoxyribonucleic acid
EBI	European Bioinformatics Institute
ENA	European Nucleotide Archive
Kb	Kilobase pairs (1000 base pairs)
M	Molarity (moles per litre)
µl	Microlitre(s)
ml	Millilitres(s)
NCBI	National Center for Biotechnology Information
OTU	Operational Taxonomic Unit
p.p.m.v.	Parts Per Million by Volume
PCR	Polymerase Chain Reaction
RDP	Ribosomal Database Project

Chapter 1: Literature Review

This literature review explores scientific literature related to extremophile microorganism growth and persistence using select volcanic gases; carbon monoxide (CO) and hydrogen (H_2). Section 1.1 covers volcanism, extreme ecosystems and the volcanic systems in in the Tongariro National Park while Section 1.2 delves into extremophilic microorganisms, microbial diversity associated with volcanic systems. Section 1.3 covers the sources and sinks of CO and H_2 and the enzymes required for trace gas (CO and H_2) utilisation. Section 1.6 defines the hypothesis and Section 1.7 defines the aims and objectives of this thesis.

1.1 Extreme and Volcanic Ecosystems

1.1.1 Extreme Ecosystems and Extremophiles

Ecosystems are generally considered extreme if their physicochemical, environmental or geographical parameters are restrictive for survival of most upper life forms[7]. The designation "extremophile" is a term to describe organisms capable of life within extreme ecosystems. Extremophilic taxa are dispersed in all three domains of life yet are predominantly prokaryotic microorganisms but also include protists (e.g., algae, fungi and protozoa) and other multicellular organisms[8]. The criterion for an extremophile is the conditions for growth fall outside the narrow parameters required for the survival of humans and most metazoan species (20-35 °C and circumneutral pH). This definition is based around the anthropogenic limits for life, therefore what extremes are from a microbial perspective remains questionable[9]. These extremophilic organisms tend to be unicellular and can be further divided into those who are tolerant of extreme environments (extremotolerant) and those which require extreme conditions for growth and multiplication, which are the two critical parameters necessary for competition in any environment. No matter the definition for extremophiles, one thing is certain: microorganisms persist and thrive in some of the harshest environments on Earth[9]. At the crux, the aim of studying extremophiles is to understand what the upper limits and features of life are and to answer the question:

are these microorganisms active, thriving, or persisting in extreme environments, or merely dead cells that were propitiously deposited by atmospheric processes?[10].

Extreme environments cover many terrestrial, marine and aquatic extremes and are determined extreme by physicochemical or atmospheric parameters such as aridity, hyper-salinity, acidity or alkalinity, nutrient limitation, permafrost, elevation (low pressure, elevation gradients), intense UV-radiation, high winds, eruptive events and intense volcanic/geothermal activity[11]. These extreme ecosystems include many anthropogenic and natural biotopes such as geothermal or hydrothermal areas, volcanic soils, the Arctic and Antarctic, non-polar hot and cold deserts, salt plains, the cold and dark ocean, the deep biosphere, submarine hydrothermal vents, wastewater treatment plants, UV-heated soil, the phyllosphere (aerial surfaces of plants), acid mines and deep dark cave systems[9,12-18]. Extreme ecosystems and extremophiles have become model systems for astrobiologists to study the likelihood and mode of life on other planets in our solar system[10].

1.1.2 Geothermal and Volcanic Geological Theory

Volcanism is the manifestation of internal thermal processes through the emission at the surface of solid, liquid or gaseous products[19] and occur on planets that have yet to cool, are comprised of tectonic plates and contain magmatic centres[20]. Earth's crust is comprised of 17 major rigid tectonic plates that float on a hotter, softer layer on its molten mantle. This mantle contains minerals, predominantly metamorphic rock, e.g. xenoliths, which can have fluid inclusions with thousands of parts per million by volume (p.p.m.v.) of volatile chemicals and acts as a reservoir for movement of chemicals expressed due to tectonic plate activity, including subduction zones and back-arc basins permit the movement of chemicals including water (H_2O), carbon dioxide (CO_2) and nitrogen (N) and sulphur (S) from Earth's interior to exterior, and vice versa. Volcanism inextricably supports life through the injection of these volatile compounds into the troposphere[20]. Environments resulting from volcanism manifest as many surface expressions including geothermal hot springs, geysers, fumaroles, deep-sea hydrothermal vents, to volcanic rock and soil environments[21]. The most abundant product of volcanism is volcanic soils and rock, which are often extreme and xeric habitats and harsh habitats even from the perspective of microorganisms[20]. Passive

volcanic degassing releases water vapour, sulphur dioxide (SO_2), hydrogen sulphide (H_2S), carbon dioxide (CO_2), hydrogen (H_2), oxygen (O_2), nitrogen (N_2), carbon monoxide (CO), methane (CH_4), argon (Ar) and helium (He), as well as other trace components into the atmosphere[22]. The composition of gases released due to volcanic degassing is the result of deep processes, such as vapour-melt separation throughout the rise of the magmas and shallow processes in the volcanic structures[22].

Geothermal environments are associated with thermal signatures which occur where crystallisation of magmatic intrusions in the crust provides a source of heat to the surface, resulting in geothermal systems where hot upflow reaches the crustal surface, creating a series of geothermal features. The source of hydrothermal fluid in terrestrial systems has been determined by stable isotope probing to be mostly from meteoric sources[23]. Other sources of hydrothermal fluid have been determined to be seawater or a hybrid mixture between seawater and meteoric sources, especially in coastal regions and oceanic islands[24]. Hydrothermal waters are primarily characterised by their aqueous geochemistry, typically, the concentration of three dominant anions: chloride, sulphate and bicarbonate. The geochemistry of surface features is largely controlled by temperature and the presence of other sources of groundwater fluids[21].

Hydrothermal systems occur when this groundwater percolates through permeable crustal rock to the level of magmatic intrusion, which can occur at depths of more than 5 km. In these regions, thermal convection is the dominant behaviour of groundwater in the permeable portion of Earth's crust where groundwater meets cooling magmatic systems, and upflow of hydrothermal fluid reaches the surface[25]. Heated groundwater fluid interacting with magmatic intrusions can reach temperatures of more than 200 °C without boiling[23]. As this heated fluid returns to the surface, the decrease of pressure can lead to boiling. When boiling, the fluid separates into aqueous and gaseous phases. If the aqueous phase reaches the surface, it can be expressed as a hot spring, geyser or boiling mud pool. The surface manifestation of the gaseous phase are fumarole gas vents, or steam, which can be absorbed by superficial groundwater [23]. Geothermal systems are often associated with caldera structures, in the case of Rotorua and Okataina Calderas in New Zealand, Yellowstone Valley in the U.S.A. and Krafla in Iceland[23].

1.1.3 Mt. Tongariro and the Taupō Volcanic Zone

The North Island of New Zealand (Te Ika a Māui, Aotearoa) is home to the Taupō Volcanic Zone (TVZ) an active, mostly rhyolitic volcanic zone with many geothermal expressions.

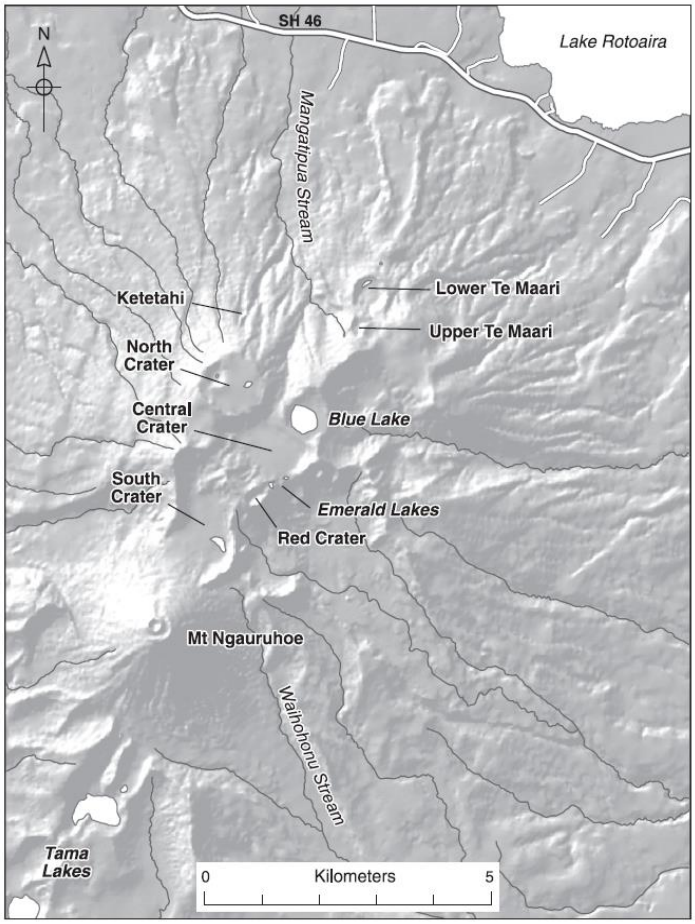

Figure 1: Map of Mt. Tongariro stratocone complex showing location of craters and vents mentioned in text. *Map retrieved from Scott el al., 2014*[6].

The TVZ extends 260 km from White Island (Whakaari) in the north-east to Mt. Ruapehu in the south-west of the Central Plateau and represents the southernmost portion of the ~ 2800 km long Tonga-Kermadec arc system[24,26]. The TVZ is an active rifting arc which encompasses 23 known hydrothermal systems, characterised by intense active crustal magmatism and rhyolitic volcanism[21,26]. Arc-type volcanism occurs when two crustal plates collide against each other along convergent tectonic plate boundaries which are characterised by small (<0.1 km^3)

eruptive events[27]. Geophysical data suggest that magma is still present below most of the TVZ, meaning stress and faulting related with rifting events could trigger, or control, eruptive events[28]. TVZ is a region of intense volcanism, associated with the subduction of the Pacific Plate beneath the crustal plate of Zealandia, on which the North Island resides[21,26,29] and supports a multitude of surface hydrothermal expressions including hot springs, geysers, mud pools, fumaroles, crater lakes, and active volcanoes[21].

Mt. Tongariro is part of the TgVC (Tongariro Volcanic Centre), which is positioned adjacent to Mt. Ruapehu on the southernmost point of the TVZ[30,31]. TgVC includes four andesite (composite of basalt and rhyolite) volcanoes –Mt. Tongariro, Mt. Ruapehu, Kakaramea and Pihanga. Mt. Tongariro itself is a stratocone complex with 17 nested and overlapping vents covering an area 5 km long by 13 km wide[27]. Mt. Tongariro consists of South, Red, Central and North Craters and two active fumarolic vents: Upper Te Maari and Lower Te Maari (Figure. 1). Historically, Mt. Tongariro has erupted diverse lava flows, including andesites, basaltic andesites and dacite with striking differences in composition between different eruption events, which has allowed detection of separate eruption events over time[27]. Underlying Mt. Tongariro is a large vapour-dominated system that is overlaid by a thick condensate layer[32]. Surface manifestations at Red Crater, Central Crater, Lower Te Maari craters and Ketetahi Hot Springs are linked to this geothermal system. No known eruptions of Mt. Tongariro have occurred since the inception of the GeoNet project in the 1940's except the 2012 eruptions at Te Maari – see below[33]. However, prior to 1940, up to fifteen eruption events are thought to have occurred between 1840-1928 based upon geological and historical accounts[6].

Unrest at Mt. Tongariro and Mt. Ngauruhoe has been recorded between 2001-2012. In 2006, shallow earthquakes were frequent beneath the South Crater area of Mt. Tongariro and were attributed to unrest at Ngauruhoe[34]. Following July 2012, the degree of seismicity has increased beneath northern Tongariro. Geochemical sampling of fumaroles revealed magmatic signatures indicating activity of fumaroles near Te Maari crater, but not at the other crater sites of Mt. Tongariro[6]. Despite these findings, Red Crater, a fumarolic surface expression within the Tongariro Volcanic Zone (TgVC), appears to be more active than previously anticipated with over seven eruption events between 1855-1934[6]. Observed unrest has also been uncovered at South Crater, with a fumarole reported early in the 19th

century, and shallow earthquakes detected by GeoNet monitoring between 2006 and 2008[6].

1.1.4 Microbial Ecology of Volcanic Rock and Soil, and Oligotrophic Ecosystems

From the perspective of a prokaryotic microorganism, environmental extremes are regarded from one major viewpoint: the availability of liquid water[2,35]. Thus, the paucity of water in hot deserts, cold deserts, high-alpine and volcanic soils or rock were once seen as life-limiting environments. Indeed, extreme high-alpine and volcanic environments are associated with a shift in biodiversity from a macro to a micro scale and a significant decrease in macro-and microbial abundance[36] and a shift towards communities with simplified phylogenetic diversity[37]. However, life is still abundant even in these remote and tough environments, despite being oligotrophic (low in nutrients, C and N)[1] and lacking eukaryotic primary producers; due to lack of water availability below the calculated 1.3% moisture limit for eukaryotic life[38]. It is worth mentioning there has been no set moisture limit for prokaryotic life, which explains why arid systems are over represented by bacteria and archaea[38,39]. These environments can have extreme fluxes in physical and chemical parameters, especially in seasonality, from high rainfall to low rainfall, high UV radiation, hot to cold temperature (sometimes covered in permafrost), high wind velocities, elevation and its associated thin atmosphere, low nutrient availability and geographic isolation[38]. An interesting question is how do microbes persist in these oligotrophic environments?

One of the most widely studied group of microorganisms that cope in these extremely oligotrophic conditions are from the phylum Actinobacteria. In studies of Icelandic basaltic volcanic deposits, it was found that representatives from the phylum Actinobacteria (genus *Streptomyces*) can grow on rock-derived nutrients and produce acid which releases elements from the volcanic rock[40]. In corroboration, Actinobacteria from the order *Actinomycetes* isolated from granite rock minerals in St. Katherine, Egypt which revealed that 97% of the detected Actinobacterial rock colonisers can produce acids to release Cu, Fe, Cd, Ag and Zn from the rocks[41]. Actinobacteria, which also sporulate, have been demonstrated to

be a pioneering rock colonising terrestrial-phylum which are strongly associated with volcanic rock weathering, which is postulated as an evolutionarily protracted role based on their heritage. They are therefore of interest as they establish life on planetary crust and due to their resilient nature, they have been of interest to astrobiologists seeking to understand how life on other plants in Earth's solar system could sustain life[40]. Actinobacteria are a globally omnipresent, ancient terrestrial lineage which are postulated to be prevalent due to their ability to colonise rocks and oxidise trace gas including CO and H_2, forming an alternative source of primary production on which communities can establish[37,40]. Actinobacteria is postulated to be one of the first terrestrial phyla, and it resides in the Terrabacteria superphylum. Actinobacteria also display tolerance to desiccation due to their Gram-positive peptidoglycan layer. Additional members of the Terrabacteria superphylum, including Cyanobacteria, Actinobacteria, Firmicutes, Chloroflexi and Deinococcus-Thermus, contain resistance to extreme environments (e.g., desiccation, ultraviolet radiation, and high salinity) and oxygenic photosynthesis, which allow these taxa to inhabit and persist on terrestrial surfaces[42].

Arguably, all microbial populations active in volcanic, high-elevation alpine environments are among the niftiest on Earth. If the environment lacks conventional primary producers (phototrophs: oxygenic photosynthetic algae, vascular plants, cyanobacteria) and input of organic matter is low, then input of carbon and energy into the system must be sourced by other avenues. Many studies have shown metagenomic, *in vivo* and *in situ* evidence of trace-gas oxidation, which collectively suggest primary production may be obtained by taxa which oxidise atmospheric trace gases CO_2, CO and H_2[1-3,37,43,44]. Microorganisms in soil microcosms have been shown to scavenge atmospheric trace gases to persist in the hostile oligotrophic Antarctic Dry Valleys[1]. These gases are also shown to serve as dependable sources of energy and carbon for thermophilic and heterotrophic microorganisms in times of carbon limitation[44,45]. In addition, CO has been shown to be a source of supplemental energy to carboxydovores, a group of mixotrophic CO-oxidisers, which oxidise CO to gain electrons, coupled to the reduction of organic matter for carbon[46]. Carboxydovores are more environmentally relevant than carboxydotrophs as they uptake CO at atmospheric concentrations, whereas

carboxydotrophs require >1% CO (10,000 ppm) partial pressure for the carbon monoxide dehydrogenase enzymes to be active[46].

To a certain extent, at the resolution of phylum level, most terrestrial habitats across earth harbour the common terrestrial phyla: Proteobacteria, Acidobacteria, Actinobacteria, Verrucomicrobia, Bacteroidetes, Chloroflexi, Planctomycetes, Gemmatimonadetes, and Firmicutes[47,48]. However, arid hot or cold deserts, volcanic soils and high-elevation alpine bacteria harbour some differences in their phylogenetic diversity compared to the average soil community. A 2010 biogeographical study of high-alpine bacteria showed that spatial distance between microbial communities less accurately predicted community assembly compared to biogeochemical and habitat parameters. It explains that habitat predicts which phyla will be dominant in a system[49]. For instance, a study of Socompa Volcano in the Atacama Desert showed that bacterial and archaeal diversity in high elevation volcanic soils depends entirely on physicochemistry (fumarolic gas input, temperature, C, soil moisture). Comparing non-fumarole, warm-fumarole and cold-fumarole soils, they found the cold-fumarole soils (-5 °C) were above 10% average soil moisture and had similar physicochemistry to low wetland soils. These cold-fumarole soils were fundamentally different from the non-fumarole (oligotrophic, arid) and warm-fumarole soil types in community composition, containing Acidobacteria (28%), Alphaproteobacteria (13%), Cyanobacteria (11%), and Chloroflexi (9%) and candidate phyla GAL15, SPAM, WPS-2, AD3, D-T and OP-10 (Armatimonadetes)[2]. Soil communities hyperarid environments tend to be 'simpler' communities, devoid of phototrophs due to moisture limitation. Furthermore, Robinson Ridge soils in the Antarctic Dry Valleys contains low-abundance of phototrophic Cyanobacteria (0.28% relative abundance), instead the communities are dominated by Actinobacteria, Chloroflexi, Proteobacteria, Acidobacteria, and two candidate phyla, AD3 (*Candidatus* Dormibacteraeota) and WPS-2 (*Candidatus* Eremiobacteraeota)[1]. Analysis showed that the candidate phyla AD3 and WPS-2 are sister lineages to Chloroflexi and Armatimonadetes[50] and are both affiliated with the Terrabacteria superphylum[51]. They are also postulated to be trace-gas scavengers of H_2, CO and CO_2 and are suggested to serve as the primary producers in oligotrophic environments devoid of traditional primary production due to environmental limitations[1]. As mentioned above,

studies of arid soils and volcanic systems all return comparable results: dominated by Actinobacteria or Chloroflexi and devoid of detectable phototrophs and have the capability to oxidise trace gas[1,37]. In contrast, when there is moisture in the system, there are detectable phototrophs in volcanic soils which shifts the primary production from alternate supplemental energy forms such as trace gas oxidation to photosynthetic processes[2].

1.2 Extremophiles

1.2.1 Extremophile Diversity and Growth

Extremophiles are microorganisms which thrive in extreme ecosystems (see Section 1.1.1.). Often extreme ecosystems are associated with reduced microbial diversity[52]. The phylogenetic diversity of extremophiles is broad and complex, particularly as some genera contain almost exclusively extreme taxa; whilst other genera contain both non-extremophilic and extremophilic taxa[8]. In addition to this, physiologically or functionally similar extremophiles are dispersed widely in the phylogenetic tree of life; present in all three domains. Extremophile is a blanket "othering" term to categorise any taxa which thrive outside the "normal" limits of life. Despite these limitations, many extremophiles have been isolated and detected in many ecosystems worldwide. Extremophiles are categorised primarily by their growth conditions (see Section 1.1.)

Extremophiles are predominantly prokaryotic microorganisms, but also include numerous protists (e.g., algae, fungi and protozoa) and other multicellular organisms[8,53]. There are diverse ecotypes of extremophiles; those which thrive in elevated temperatures are termed thermophiles or hyperthermophiles and conversely, those which enjoy extremely low temperatures are psychrophiles. Barophiles thrive in extreme pressure systems such as the deep-sea whilst halophiles thrive in extreme concentrations of salt. Acidophiles thrive in acidic pH environments whilst alkaliphiles require alkaline pH surroundings[7]. Other extremophilic taxa live in areas of heavy metals (metallophiles), large doses of radioactivity or in oligotrophic environments. In contrast, organisms which thrive in slightly warmer than average or ambient environments they are termed mesophiles or neutrophiles, respectively[54].

If an organism has adapted to thrive in multiple extreme conditions, it's a polyextremophile. For example, a thermoacidophile is adapted to both hot and acidic conditions[8]. Hydrothermal zones associated with volcanism have surface features which reach surface boiling point temperatures (100 °C, dependant on altitude) are well known to support both thermotolerant, thermophilic and hyperthermophilic organisms. The classification of thermophilic can be broadly sub-classified into three groups based on their optimal growth temperatures; moderate thermophiles (45-70 °C), thermophiles (70-80 °C) and hyperthermophiles (≥80 °C)[9]. Most thermophiles are capable of slow growth between 25- 40 °C[55]. Thermophily is an adaptation spanning diverse environments and taxonomic groups which have different functional and metabolic capabilities, dependant on the species.

Hyperthermophilic growth temperatures represent the upper temperature for life, with growth between 60-121°C[56] and an optimal temperature range being 80-110 °C[57]. Hyperthermophiles are found almost exclusively in either geothermally or volcanically heated biotopes, many situated along tectonic spreading and subduction zones either terrestrially or in submarine zones. The fundamental prerequisite for life is water, more specifically, the availability of liquid water. So, when the boiling point of water increases in zones of elevated atmospheric, hydrostatic or osmotic pressure, this leads to hyperthermophily in which growth temperatures can exceed 100 °C[57], hence why hyperthermophiles are archetypally isolated in oxygen-free terrestrial and submarine biotopes[57].

Protozoa, fungi, algae, and many taxa from bacterial and archaeal phyla share thermophily. Examples of phyla which contain thermophilic species include Deinococcus-thermus (*Thermus*), Aquificae, Thermotogae, Thermodesulfobacteria, Chloroflexi, Nitrospirae, Deferribacteres, Proteobacteria (*Thiobacillus*), Firmicutes (*Bacillus, Clostridium, Desulfotomaculum*), Actinobacteria (*Actinomycetes*), Spirochaetes, Bacteroidetes, Caldiserica, Armatimonadetes (class Chthomomonadetes), Dictyoglomi, Spirochaetae, Cyanobacteria, among many more[58,59]. Thermophilic Archaea include: *Pyrococcus*, *Thermococcus*, *Thermoplasma* and *Sulfolobus*[7]. Hyperthermophilic taxa are predominantly found within the domain Archaea in the major phyla Crenarchaeota (*Pyrolobus fumarii*), Euryarchaeota and Nanoarchaeota, with the

exception the bacterium, *Thermotoga maritima* in the phylum Thermotogae, and select taxa from the phylum Aquificae[57,60].

Environments for moderate thermophiles span from wastewater sludge to sun-heated soils, hydrothermal water to volcanic soils. Thermophiles and hyperthermophiles generally reside in terrestrial and marine geothermally heated environments. A rule of thumb for hyperthermophiles is their inability to grow, even slowly, under 60 °C[55]. Acidophiles are found predominantly within the bacterial phyla Verrucomicrobia, Acidobacteria, Firmicutes, Nitrospirae, Alphaproteobacteria, Betaroteobacteria, Epsilonproteobacteria, Aquificae, Actinobacteria, and in the archaeal phyla Crenarchaeota and Euryarchaeota[61]. *Cyanidium caldarium* is a red alga capable of growth at pH 0.7 and 60 °C[7]. Alkaliphiles consist of alkaliphilic and haloalkaliphilic organisms, which require a pH of 9 or above. Aerobic alkaliphiles are found in the phyla *Bacillus*, *Micrococcus*, *Pseudomonas*, and *Streptomyces*. Alkaliphilic eukaryotes including yeasts and filamentous fungi have also been studied. Anaerobic alkaliphile isolates include *Clostridium thermohydrosulfuricum* 39E, *Anaerobranca horikoshi*, *Clostridium paradoxum*, *Thermococcus acidaminovorans*[55]. Barophiles are mainly represented in Gammaproteobacteria[55]. Radiation resistance is represented in the bacterial species *D. radiodurans*, *Rubrobacter* and the green alga *Dunaliella bardawil*[7].

There is one commonality with most known extremophilic taxa: they reside in common phyla, meaning phylogeny gives little insight into their physiological capability. This highlights the importance of culturing extremophiles in addition to 16s rRNA gene sequence community profiling or metagenomics.

1.2.2. Survival and Persistence of Extremophiles

Very little research has focused on how extremophiles survive outside their normal growth environments. However, it is well known that bacteria within oxic environments often exist as dormant cells, particularly in times of organic carbon limitation. Bacterial cells persist by decreasing metabolic output and focusing on basal metabolism and sourcing alternative energy forms[45]. There are well known mechanisms for survival and persistence documented for general microorganisms,

including sporulation, dormancy ("microbial seed banks"), bet-hedging and resource stockpiling e.g. glycogen or polyhydroxyalkanoates[62-64]. Insights into persistence have been illuminated through recent publications of oligotrophic environments which contain very few traditional primary producers. In Antarctic Dry Valley soils for example, microorganisms survive in various states of dormancy, with the majority of their metabolic income being spent on cell maintenance, rather than growth[1]. The primary producers in oligotrophic soils are commonly cyanobacteria and algae (phototrophs), yet in Robinson Ridge, a dry valley in Antarctica, phototrophs are only present in low abundance, highlighting that these communities must be obtaining alternative sources of energy[1]. Indeed, based on metagenomic surveying of these soils the dominant genes for metabolism were for CO_2 fixation via the Calvin-Benson-Bassham cycle - genes were detected for type IE ribulose-1,5-bisphosphate carboxylase (RuBisCO). This RuBisCO lineage is known to support hydrogenotrophic growth of Actinobacteria[1]. High affinity enzymes responsible for aerobic respiration of molecular H_2 and CO were also present. The lineages identified are group 1h [NiFe]-hydrogenases (structural genes: hhyl, hhys) and type 1 [MoCu]-carbon monoxide dehydrogenases (structural genes: coxL, coxS and coxM); these two lineages are known to support trace-level atmospheric scavenging of H_2 and CO respectively[65,66]. These metagenomic predictions were confirmed by microcosm experiments where Robinson Ridge soils consumed headspace CO and H_2 to below tropospheric concentrations confirming high affinity hydrogenases and CO dehydrogenases were present, which was corroborated by using PCR and primers to target the large subunits of the genes responsible for the production of these enzymes[1].

There are multiple studies of ecosystems and isolate cultivation experiments that support the observations that dormant heterotrophs, such as those in Antarctica's Dry Valleys are capable of energy generation by oxidation of trace levels of H_2 and CO[1]. For example, a metagenomic study of trace gas oxidation in the hyperarid Atacama Desert has been observed and linked to Actinobacteria which encode genes required for scavenging trace gas[37]. Similar observations have been made for microbial populations inhabiting recent volcanic deposits in Hawaii which depend on atmospheric CO and H_2 as sources of energy and carbon in areas lacking vegetation[3]. Upon further inspection, the genes responsible for hydrogen and carbon monoxide oxidation seem to be fairly pervasive in soil environments,

meaning they could be an important source of persistence for soil communities in times of nutrient-limitation[1,44,46,66].

In support of these ecosystem observations, cultivation-based experiments show that a heterotrophic microbial strain can persist under carbon-limitation via trace gas oxidation. For example, the acidobacterial moderately thermophilic isolate (*Pyrinomonas methylaliphatogenes* K22^T) from the volcanic soils of Mt. Ngauruhoe (New Zealand) (Figure 1) can scavenge trace levels of H_2 under nutrient limitation using a respiratory-type 1h [NiFe]-hydrogenase enzyme. The genes responsible for this enzyme were upregulated in stationery phase, linking them to persistence and dormancy. This highly ubiquitous but understudied phylum, Acidobacteria, have a *K*-strategist slow-growth approach with a tendency towards dormancy[44]. Similar observations exist for two moderately thermophilic Chloroflexi isolates, *Thermomicrobium roseum* (class Chloroflexia) and *Thermogemmatispora* sp. T81 (class Ktedonobacteria), which also upregulate genes for type 1h [NiFe]-hydrogenases and type 1 [MoCu]-carbon monoxide dehydrogenases in order to persist in carbon-limiting conditions[45].

Hydrogen and CO have been revealed as dependable atmospheric sources of energy and carbon for cells in dormancy due to their ability as reductants and ubiquity of H_2 and CO as trace gases in the troposphere[1,4,37]. The broad distribution of trace atmospheric gas utilisation prompts the question: **are trace gases a source of primary production and microbial persistence in all oligotrophic systems?**

1.3. Atmospheric Trace Gases

Geothermal and volcanic gases are specific to individual sites and are a result of passive magmatic degassing and gas phase separation, as well as shallow interactions with near surface geology and groundwaters[22]. While a variety of gases are released from volcanic systems (e.g. water vapour, SO_2, H_2S, CO_2, H_2, O_2, N_2 CO, CH_4, Ar and He; see Section 1.1.2), this section of the literature review focusses exclusively on the sources and sinks of CO and H_2 gases and the enzymology of trace gas oxidation.

1.3.1 Sources of Carbon CO

Carbon monoxide is produced by natural abiotic, biotic and anthropogenic processes. Production of CO in the northern hemisphere is two-fold greater than the southern hemisphere[67]. The annual production of CO is approximately 2,500 Tg[68]. Major sources of anthropogenic CO include incomplete combustion of fossil fuels (accounts for 15% of total budget), biomass burning and the oxidation of biosphere hydrocarbons[67]. Atmospheric sources include the oxidation of natural and anthropogenic CH_4 and non-methyl hydrocarbon (NMHC) sources (25%)[67,68]. Open-ocean photolysis of dissolved organic matter is the dominant source of marine CO production, though the ocean contributes only 10% of all anthropogenic sources of CO[69]. Volcanic outgassing also releases low volumes of CO and are commonly detected via satellite observations[70]. The biospheric sources of CO production include forest canopy emissions, root nodules, photo and thermochemical degradation of organic matter in marine sediments and terrestrial soils[46]. Carbon monoxide is also produced by anaerobic bacteria living in anoxic zones of aquatic environments[71] and oxidation of CH_4 to CO by methanotrophs[72].

Carbon monoxide is an extremely versatile and significant molecule on Earth, despite holding the reputation of a noxious gas. It has contributed to many important processes in the history of life across scales that range from cellular to planetary and has many influences on the composition of the atmosphere. Indeed, numerous studies indicate CO is involved in more reactions than any other biologically relevant molecule[46]. Prebiotic chemistry models place CO as a vital reactant in numerous key syntheses[46,73-75].

Concentrations of CO in the early Earth atmosphere are estimated to be in the range of 100 ppm[46] but decreased to current levels of between 50-350 ppb (0.05-0.35 ppm) over geological time[76]. The molecule has a relatively short lifetime of 10-52 days[77] and despite CO being at trace levels and absorbing very little infrared radiation, it affects the oxidative capacity of the lower atmosphere due to photochemical oxidation of CO by hydroxyl radicals (OH), release of ozone (O_3) and direct radiative forcing[76]. Up to 70% of annual CO production is oxidised by OH[78]. Carbon monoxide can affect global radiative forcing directly by absorbing and emitting infrared wavelengths between 1800 and 2300 cm^{-1}, and indirectly by altering the amounts of other radiatively active gases including CH_4, O_3 and carbon dioxide (CO_2). Though atmospheric levels of methane exceed those of CO by up to 17-fold, emission of CO (about 2,500 Tg $year^{-1}$) surpasses the emissions of methane by around fivefold. This incongruity is due to the rapid reaction between CO and OH[76].

The OH ion is a volatile molecule with a very short lifetime (about 1 s) due to the high reactivity of unpaired electrons which tend to favour transfer of electrons to other molecules[79]. CO and CH_4 are the two major direct sinks of atmospheric OH. Globally, the oxidative reaction between OH and CO is about twice the rate of OH with CH_4, making CO the principal sink of OH[80]. Conceivably, if levels of trace gases increase, this results in depleted levels of OH left in the atmosphere to react with major greenhouse gases such as CH_4 and carbon dioxide (CO_2). In effect, CO acts as an indirect greenhouse gas through its swift reaction with OH which leads to increased decay times for CH_4 and CO_2[81]. Indeed, some predictions estimate that CO emissions will exceed emissions of N_2O due to anthropogenic activity in the short term[82]. Finally, numerous studies suggest that injection of CO into a chemical system containing adequate nitric oxide (NO) results in the photochemical production of O_3, whilst emissions of CO can lead to O_3 loss in regions with low levels of NO (<10 ppt)[82,83].

1.3.2 Sources of H_2

Molecular hydrogen is an indirect greenhouse gas which is present at trace levels in Earth's lower atmosphere (stratosphere) and has a rapid turnover rate (lifetime $\leq$ 2.1 y^{-1})[84]. It is the second most abundant reduced gas in Earth's troposphere after CH_4.

The sum of H_2 sources and sinks are presently near equilibrium; H_2 maintains the current atmospheric oxidation state due to involvement in complex biogeochemical reactions. It is well established that anthropogenic and geochemical processes (including volcanic outgassing) represent the majority of net H_2 production[85]. The largest atmospheric sources of H_2 are from CH_4 oxidation and photochemical oxidation of non-methyl hydrocarbons NMHC including ethylene and terpenes. Methane is a direct radiative forcing greenhouse gas emitted predominantly from wetlands and the terminal product of methanogenesis by anaerobic archaea[86]. In urban areas, measurements suggest fossil fuel burning is the main anthropogenic source of tropospheric H_2. In corroboration, measurements show mixing ratios of H_2 in urban areas follow a bimodal diurnal trend, correlating with peak morning and afternoon traffic[85]. Hydrogen is consumed and produced by ecosystems spanning from deep sea hydrothermal vents, rood nodules of plants to volcanic soils, sourcing energy to aerobes, anaerobes, methanogens, acetogens, sulphate reducers, phototrophs and many more metabolic types[3,87-89]. Other sources of H_2 include biomass burning, phosphite (HPO_3^{2-}) oxidation, nitrogen fixation in root nodules, H_2 supersaturation in surface oceans (especially in biomass rich zones). If supersaturation of H_2 in surface oceans (the pelagic zone) is uniform, estimates suggest that around 6% of global H_2 emissions could be contributed by the pelagic zone alone[90]. Though H_2 is produced in root nodules by microbial nitrogen fixation, it is quickly consumed by low-affinity H_2-oxidising bacteria within 5 cm from the nodule surface in the humus layer, thus not contributing significantly to tropospheric H_2 levels[85]. Nitrogen fixation in seas and lakes is a significant source of atmospheric H_2. Submarine hydrothermal vents also release H_2-rich hydrothermal fluids, where terrestrial life is hypothesised to have originated[91]. Also, H_2 is released as a terminal product from metabolism in facultative and obligate fermenters.

In fact, H_2 is known as the "universal energy currency" in anoxic environments, connecting diverse physiological taxa caught in a state of flux between consumption and production of H_2 in a complex food web, which is often referred to as interspecies H_2 transfer [87].

1.3.3 Trace Gases Introduction

Earth's atmosphere is the result of dynamic chemical (abiotic) and biological processes. Microorganisms, particularly soil microorganisms, contribute to the atmospheric budget of many trace gases by acting as sources and sinks of biologically relevant gases[92]. Gases which have a mixing ratio of < 1% (10,000 ppm) in the atmosphere are considered trace gases. The most important of these gases influenced by soils include CH_4, CO, OCS (carbonyl sulphide), H_2, N_2O and NO[93]. The relevance of trace gases in Earth's lower atmosphere (troposphere) on atmospheric chemistry and climate change pertain to their impact on the oxidative capacity of the atmosphere through their impact on the stability of aerosols, ozone (O_3) production and destruction, greenhouse and pollutant gases[93]. Trace gases such as H_2 and CO react with OH radicals, indirectly increasing the concentration of direct greenhouse gases such as CH_4 which are usually removed by OH, the atmospheric detergent. Therefore, the flux of trace gases plays a key role in controlling atmospheric chemistry, greenhouse gas concentration and consequentially, climate change[79].

The role microorganisms play in tropospheric cycling of trace gases is well studied for CH_4 and inorganic gases including nitrogen oxides and ozone. Trace-gas oxidation of molecular H_2, CO and many volatile organic compounds of biogenic origin (bVOCs), such as monoterpenes and chloromethane[92]. Many physicochemical parameters influence rates and modes trace gas oxidation, including availability of water, sources of organic matter and amount of oxygen. This is important for respiration processes such as denitrification where the oxidation of NO_3 to N_2 is sometimes incomplete in aerobic environments, as the enzymes responsible for reducing N_2O to N_2 are oxygen sensitive. Rates for CO oxidation are also faster in environments rich in organic carbon for CO-oxidisers capable of mixotrophic metabolism where CO is an electron source for aerobic

respiration[46]. Trace-gas oxidation has broad phylogenetic, enzymatic and environmental diversity, which translates into functional redundancy in environmental microbiomes and manifests as changes in gene expression of specific taxa in several studies in times of environmental perturbation[92].

1.3.4 Microbial CO oxidation

Oxidation of CO was likely favoured in Earth's early history due to CO levels which are theorised to have exceeded 100 ppm [75], predating the formation of an oxidising atmosphere (due to the inception of oxygenic photosynthesis)[94]. Numerous studies reinforce the hypothesis that hydrothermal vent communities survived using CO as their sole energy and carbon source, using the extant species *Carboxydothermus hydrogenoformans* as a model for primitive CO oxidising chemolithotrophs. Extant hydrothermal vents harbour populations of anaerobic CO oxidisers and the addition of oxygenic photosynthesis into the blend generated oxidants such as oxygen, nitrite, nitrate and sulfate which extended the use of CO as an energy source[46,88,95,96,97]. Aerobic CO oxidation is a ubiquitous metabolism which is utilised by phylogenetically and physiologically diverse archaeal and bacterial microorganisms which inhabit most ecosystems on Earth; including aquatic and terrestrial zones[46,94]. CO has a short lifetime in soil due to aerobic CO oxidation occurring primarily in the top few centimetres of the humus layer (O horizon)[98]. Physiological and phylogenetic diversity of bacteria and archaea which utilize CO as an energy or carbon source is likely due to the widespread sources of CO and the diverse habitats which support CO oxidation[46]. Atmospheric CO serves as a dependable substrate for soil microorganisms, which remove a significant portion of the CO flux annually (around 15%)[99]. Most pioneering studies of CO oxidation focus on carboxydotrophs which grow at elevated (>1%, 10,000 ppm) using CO as their sole carbon and energy source. Carboxydotrophs often couple CO oxidation to CO_2 fixation by the Calvin-Benson-Bassham cycle to generate biomass, using the enzyme ribulose-1,5-bisphosphate carboxylase (RuBisCO) encoded by the *cbbl* gene. Carbon monoxide-oxidising bacteria which lack RuBisCo often use CO solely as a supplemental energy source[46].

The majority of newly recognised CO oxidisers are inhibited by elevated levels of CO; these microorganisms are described as carboxydovores, a group of mixotrophic CO oxidisers which oxidise low levels (<1000 ppm) of CO as their energy source coupled to the reduction of organic matter to salvage carbon. Carboxydovores are considered more ecologically relevant than carboxydotrophic microorganisms as they are capable of trace atmospheric CO oxidation as a means of microbial persistence, therefore controlling the flux of trace CO in the atmosphere. Aerobic CO oxidisers couple CO oxidation to oxygen reduction using the reducing equivalents from CO oxidation. These reducing equivalents can also be utilised to reduce nitrate to nitrite (dissimilatory nitrate reduction) or nitrate to dinitrogen (denitrification)[46].

1.3.5 Enzymology of CO Oxidation

Microbial CO oxidation reaction is catalysed by the enzyme carbon monoxide dehydrogenase (CODH), which is encoded by the *cox* operon[98]. In aerobic CO oxidisers, the CODH is primarily bound to the inner aspect of the cytoplasmic membrane but can also be located in the cytoplasm[100]. This enzyme functions to either oxidise CO, synthesize acetyl-CoA, or cleave acetyl-CoA in several pathways which yield energy[101]. Anaerobic CO dehydrogenases present in methanogens and acetogens generally have Acetyl-coA activity[98]. This complex reaction is carried out by a single multimeric, bifunctional, and oxygen-sensitive five-subunit enzyme complex called carbon monoxide dehydrogenase/acetyl-CoA synthase (CODH/ACS) which exists in thousands of archaeal and bacterial genomes and is responsible for production of Acetyl-coA through the reduction of CO_2 using the Wood-Ljungdahl pathway[102]. This complex is present in both bacteria and archaea suggesting the enzyme complex originated before the last universal common ancestor (LUCA), suggesting that CODH/ACS could be conserved in genomes for over 3,500 Mya[102].

The overall reaction results in CO_2 and a pair of reducing equivalents (H^+):

$$CO + H_2O \rightarrow CO_2 + 2H^+ + 2e^- \qquad \text{Equation } 1^{103}$$

The reducing equivalents generated by CODH can be either transferred to electron transport proteins (cytochromes) or coupled to metabolic reactions by obligate anaerobes. Under anaerobic conditions, CO is utilised with low affinity by methanogenic archaea, acetogenic, sulphate-reducing, phototrophic bacteria, or specific anaerobic CO oxidisers which yield molecular hydrogen as their end product[46,98].

Generally, anaerobic CODH encode nickel-containing [4Fe-4S] proteins and can be linked to acetate metabolism. Conversely, aerobic CODH encode molybdenum-containing active sites, [2Fe-2S] proteins and flavoproteins. Aerobic CODH consist of a dimer of heterotrimers in three subunits: a 88.7-kDa molybdoprotein (large), a 30.2-kDa FAD-binding flavoprotein (medium), and a 17.8-kDa iron-sulphur flavoprotein (small), encoded by the genes in transcriptional order *coxL*, *coxM* and *coxS*, respectively[98]. The *coxL* subunit contains the active site for CO oxidation, which occurs at the molybdenum ion of a $CuSMoO_2$ cluster. The molybdoprotein carries the active site and a 1:1 molar complex of molybdopterin–cytosine dinucleotide (MCD) which coordinates the molybdenum to the active site, and a molybdenum atom[46,98].

There are two discovered forms of aerobic CODH. Form I (formerly referred to as OMP) has specificity to the oxidation of CO, whereas Form II (formerly referred to as BMS) is a putative CODH with unresolved functionality. These two forms which share between 40-50% nucleotide similarity to coxL, and the two forms both contain several amino-acid motifs, including GGFGXK, QGQHXTX, GSRST and CGTRIN[104], but differ in their active site motifs. The active site of Form I pertains a AYXCSFR motif, whereas form II contains a AYRGAGR motif. In the case of differentiating the two different active site forms, using these motifs can be useful, though some CO oxidisers contain variability in the copy number and combination of enzymes they possess and can encode RuBisCO in the case of carboxydovores. The Form II active site is unique to CO oxidisers but has strong structural homology to other molybdenum hydroxylase enzymes with a wide range of carbon substrates[105]. Their iron-sulphur protein centres, in variance between form II CODH is that the active site of form I CODH contains an essential loop of four amino acids, cysteine, serine, phenylalanine, and arginine, and a copper atom linked to the active site cysteine-sulphur and to the molybdenum atom[103].

In addition to the three structural genes of CODH: *coxS*, *coxM* and *coxL*, there are often several accessory genes present in the cox operon. The CODH of α-Proteobacterial species *Oligotropha carboxydivorans* is the best characterized example. The cox operon of *O. carboxydivorans* contains the 3 essential structural genes and an additional 9 accessory genes[106]. Three of the accessory genes, *coxD*, *coxE* and *coxF* are considered essential for CODH as they are involved with the synthesis of the $CuSMoO_2$ enzyme centre. In contrast to this, the thermophilic strain *Thermomicrobium roseum* DSM 5159 encodes form I structural subunits for type 1h [NiFe]-hydrogenase, and only requires *coxMSLF*[107]. Form I CODH are well documented in scientific literature relatively little is known about Form II. CO oxidising activity pertaining to Form II CODH has been documented only in *Bradyrhizobium japonicum* strain 110spc4 which oxidises CO as its sole energy and carbon source under aerobic and chemolithoautotrophic conditions, at extremely slow rates, indicating the enzyme has other preferential substrates[108].

1.3.6 Microbial H_2 Oxidation

Molecular H_2 is the most fundamental element, contributing approximately 75% of the universe in terms of mass[109]. The tropospheric mixing ratio of H_2 is currently around 530 ppb, making it the second most abundant oxidizable trace gas after methane (CH_4 – 1750 ppb)[110]. Molecular hydrogen is an important energy source for microbial growth and survival, serving as a major electron donor for respiration in both anoxic and oxic ecosystems[66].

H_2 is described as an indirect greenhouse gas and soils have been known to consume the gas since measurements were published in 1937[109]. There are two dominant sinks for tropospheric H_2: around 75% is removed by microbial oxidation in soils (known as dry deposition), the rest is lost by reaction of H_2 with OH (hydroxyl radicals) in the troposphere[109]. This process is extremely important in the biogeochemical cycle of H_2 and has been inspiring chemists for forty years. Contemporarily, the biogeochemical cycle of H_2 has been of great interest to microbiologists due to the H_2 cycle relevance for climate change and the prospect of H_2 as a biofuel source. The partial pressure of H_2 affects the oxidative capacity

of the atmosphere and affects the amount of water vapour available in the stratosphere[109].

The turnover of H_2 in the troposphere is estimated at a total loss of 79 Tg yr^{-1} corresponding to 39 Tmol yr^{-1} in molar units[109]. Microbial H_2 oxidation has been a known process since 1906, with the discovery of Knallgas bacteria capable of hydrogen oxidation in microaerophilic conditions in root nodules[111,112]. Most of the identified H_2-oxidising bacteria (HOB) are capable of H_2 metabolism at very high concentrations produced in geothermal or biological processes, i.e. their threshold for H_2 far exceeds tropospheric concentrations of the gas[99].

However, H_2 produced by biological processes (fermentation, symbiont nitrogen-fixation in root nodules) generally have no net effect on tropospheric concentrations, as these sources of H_2 are largely recycled by intraspecies and interspecies H_2 transfer without escaping to the atmosphere. Therefore, the dominant influence on atmospheric H_2 is trace-gas scavenging by soil microorganisms [44]. The first microbial isolate to demonstrate high-affinity tropospheric oxidation of H_2 was *Streptomyces* sp. PCB7 in 2008[113].

Microbial trace-gas oxidation is one facet of the H_2 cycle which has the most importance for atmospheric flux of H_2, yet the process is has only been studied over the past few years in detail[44]. Recently, evidence has shown that aerobic soil isolates from the phyla Actinobacteria and Acidobacteria scavenge H_2 from the troposphere, overturning the widely held axiom that hydrogen oxidation is restricted to low O_2, high H_2 environments[44,84,114]. Hydrogenases are widely distributed in the domains bacteria and archaea. In Eukaryotes, hydrogenases were discovered in unicellular algae [115], hydrogenosomes of unicellular protozoa[116] and recently, in hydrogenosomes-like organelles in the metazoan phylum *Loricifera*. *Loricifera* were found living in hypersaline anoxic deep L'Atalante basin which provides evidence that hydrogenases may be found in other multicellular animals[117,118].

H_2 is consumed and produced by diverse phylogenetic and functional groups, spanning environments from deep sea hydrothermal vents, rood nodules of plants to volcanic soils, sourcing energy to aerobes, ananerobes, methanogens, acetogens, sulphate reducers, phototrophs and many more metabolic types[43,87,88]. Metabolism of H_2 is observed in multiple archaeal, bacterial and eukaryotic lineages. The

diversity of hydrogenase enzymes, which catalyse the oxidation of H_2 gas, has recently been recovered in high resolution. From a metagenomic survey of many environments, diversity of phylogenetically and physiologically distinct classes of hydrogenase enzymes have been revealed[66]. Evidence suggests that hydrogenase enzymes have diversified to assume a wide range of physiological roles in ecosystems. The importance of H_2 metabolism has become increasingly apparent for a wide range of microorganisms from phototrophs, lithotrophs, respirers, fermenters to aerobes and anaerobes[66]. The diversity of hydrogenase enzymes is predicted to support H_2-based respiration, fermentation and processes linked to carbon fixation in both oxic and anoxic environments.

Furthermore, H_2 is hypothesised to be the primordial electron donor, explaining the wide range of hydrogenase enzymes and the ubiquitous, almost cosmopolitan nature of taxa capable of this metabolism, no matter the environment. In fact, there are over 51 phyla with hydrogenase-encoding genes present in their genome[66].

1.3.7 Enzymology of H_2 Oxidation

Molecular hydrogen has desirable physical properties for both macro- and microorganisms. The oxidation of H_2 is a highly exergonic process ($H_2 \rightarrow 2H^+ + 2e^-$, $E^{0'} = -0.42$ V) and the gas is easily diffusible (diffusion coefficient = 4×10^{-9} $m^2\ s^{-1}$)[87]. Microorganisms can consume H_2 as an energy source and produce H_2 with catalytic enzymes called hydrogenases[87]. Hydrogenase enzymes catalyse the oxidation of H_2 to reduce electron acceptors such are O_2, NO_3^-, fumarate, sulphate (SO_4^{2-}) or artificial electron acceptors such as methylene blue[118].

The reversible splitting of H_2 into protons and electrons is a key process in the metabolism of many microorganisms. H_2 oxidation from hydrogenase enzymes releases low-potential electrons which can be transduced through respiratory chains (to generate ATP) or to fix inorganic carbon. In reverse, H_2 is produced by the hydrogenase enzyme to expel excess reductant from the cell, in the form of diffusible gas, during microbial fermentation and photo-biological processes[87].

The overall reaction of hydrogenases results in and a pair of reducing equivalents (H^+) and two electrons:

$$\mathbf{H_2 \rightarrow 2H^+ + 2e^-} \qquad \text{Equation } 2^{44}$$

Hydrogenases are metalloenzymes which contain metal-binding motifs with [FeFe], [Fe], [NiFe] (and the [NiFeSe] subfamily) metal centres[119]. These three classes of metal centres which binds H_2 distinguish phylogenetically and physiologically unrelated classes of hydrogenases. There are four distinct groups (with 22 subgroups) of [NiFe]-hydrogenase, three groups (6 subgroups) of [FeFe]-hydrogenase, and a smaller group of [Fe]-hydrogenases[66]. Certain hydrogenases from anaerobic bacteria and archaea are also able to use protons as terminal electron acceptors in low-potential ion-translocating complexes with ferredoxin[120]. Over 3000 protein sequences for putative hydrogenases have been identified from publicly available databases[66]. According to this study, [NiFe]-hydrogenases can be divided into H_2-uptake (group 1, 2), bidirectional (group 3) and H_2 evolving clades (group 4). Upon further analysis, these clades can be split into 22 subgroups with distinct physiological roles based on phylogeny. [NiFe]-hydrogenases are the most widespread of the enzymes, present in 36 bacterial and 6 archaeal phyla. They are widespread in all classes of Proteobacteria, and Firmicutes, Cyanobacteria, Aquificae, Euryarchaeota and Crenarchaeota. [NiFe]-hydrogenases were also detected in phyla which have not been previously described including Bacteroidetes, Chlorobi, Chloroflexi, Planctomycetes and Verrucomicrobia Group 1 and 2 [NiFe]-hydrogenases which mediate respiratory H_2 uptake were encoded in 19 of the 20 ecosystems included in this metagenomic survey. They are also widely distributed in the genomes of many bacterial and archaeal species belonging to a wide range of phyla[66].

Phylogenetic analysis reveals that the deepest-branching hydrogenases belong to groups (1a, 1b) of the [NiFe]-hydrogenases. These group 1a and 1b enzymes are sensitive to O_2, mediate anaerobic respiration and occur only in strict anaerobes and found primarily in hypoxic soils, such as bog and permafrost soils. It has been hypothesised that the partial pressure of oxygen (pO_2) led to the diversification and distribution of these respiratory-type hydrogenases[66]. In the more recently branching [NiFe]-hydrogenase lineages (groups 1d, 1h and 2a) are oxygen-tolerant

enzymes which facilitate respiration in both aerobes and facultative anaerobes. These lineages were commonly detected in aerated samples such as forest and agricultural soils. Evidence suggests that these hydrogenase lineages evolved around the arrival of oxygenic photosynthesis. Furthermore, oxygen-tolerant hydrogenases are more widespread than initially thought, being present in 17 bacterial and archaeal phyla including in 9 of the most dominant soil phyla in global soils[47]. These 9 phyla are Proteobacteria, Acidobacteria, Actinobacteria, Verrucomicrobia, Bacteroidetes, Chloroflexi, Planctomycetes, Gemmatimonadetes, and Firmicutes[47]. The group which mediate tropospheric oxidation of H_2 (Group 1h [NiFe]-hydrogenases) are encoded in multiple representatives from Acidobacteria, Verrucomicrobia, Chloroflexi and Planctomycetes. Hydrogenase encoding genes are also present in multiple lineages of bacteria and archaeal methane, ammonia and nitrite oxidisers[66]. There are four distinct groups (with 22 subgroups) of [NiFe]-hydrogenase. Group 1 are membrane-bound for H_2 uptake, group 2 are cytosolic H_2 uptake, group 3 are cytosolic bidirectional, group 4 are membrane-bound H_2 evolving. The [NiFe]-hydrogenases are heterodimeric proteins consisting of closely associated small (*hhyS*), large (*hhyL*) and FeS (*hhyE*) subunits. The large subunit contains the [NiFe] centre, the catalytic site for H_2. The small subunit contains three iron-sulphur centres and depending on the group, or subgroup, the ligands of the proximal cluster (Co and CN^-) and the configuration of the medial cluster can differ, but the metal clusters remain the same. The metal cluster are configured as $[FeS]_{distal}$, $[FeS]_{medial}$, and $[FeS]_{proximal}$ clusters, i.e. the clusters are usually comprised of [4Fe4s], but they can be [3Fe4S] or [4Fe3S][66]. There are two motifs surrounding these metal-ligating cysteine residues, called L1 and L2, which can be used to distinguish between different functional groups of hydrogenase enzymes[66]. The maturation enzymes required for synthesis of [NiFe]-hydrogenases include a nickel insertase (HypA), a nickel chelator (HypB) and a putative [2Fe2S] protein (tentatively named *hhyE* – the immediate electron acceptor for the hydrogenase).

Many [NiFe]-hydrogenases are O_2-sensitive, and their catalytic activity diminishes under aerobic conditions. However, O_2-tolerant periplasmic space [NiFe]-hydrogenases can facilitate H_2 oxidation under trace atmospheric conditions. This is due to the proximal [FeS] cluster located adjacent to the enzyme active site which contains a highly reducing cluster: [4Fe3S]-6Cys, as compared to O_2-sensitive

enzymes which have a [4Fe4S] structure[121]. These enzymes can both evolve and uptake H_2, with low-potential multiheme cytochromes such as c_3, which act as either electron donors or acceptors dependent on their oxidation state. The small subunit contains three iron-sulphur clusters and the large subunit contains the [NiFe] active centre, connected to the solvent by a hydrophobic molecular tunnel[122]. In a subset of [NiFe]-hydrogenases, the Ni-bound cysteine residue is replaced by selenocysteine, these enzymes are designated as the [NiFeSe]-hydrogenase. Periplasmic, cytoplasmic and cytoplasmic membrane hydrogenases have been discovered. Generally, [NiFe] hydrogenases are biased towards oxidizing H_2, as opposed to evolving H_2. A diverse spectrum of enzymatic affinities to H_2 have also been observed in H_2-oxidizing hydrogenases[84]. Many [NiFe]-hydrogenases are O_2 tolerant[123].

The [FeFe]-hydrogenases split into three groups: fermenting and bifurcating hydrogenases (group A), an ancestral hydrogenase lineage with unknown function (group B) and putative sensory hydrogenases (group C). Group A [FeFe]-hydrogenases can be subdivided into four distinct subtypes: stand along enzymes (group A1), glutamate synthase associated hydrogenases (group A2), NADH dehydrogenases (group A3) and formate dehydrogenases (group A4). [FeFe]-hydrogenases were detected in the genomes of 12 phyla of anaerobic bacteria, 5 phyla of unicellular eukaryotes and 1 in *Diapherotrites*, a candidate archaeal phylum. Sensory and bifurcating [FeFe]-hydrogenase types are exclusive to anaerobic bacteria residing in the phyla Firmicutes, Bacteroidetes, Spirochaetes, Thermotogae and Fusobacteria. In contrast, [Fe]-hydrogenases are detected exclusively in the genomes of 25 methanogen species[66]. In addition to this, the presence of hydrogen pathways such as electron-bifurcating [FeFe]-hydrogenases can suppress methane production in ruminants[124].

There are three groups (and 6 subgroups) of [FeFe]-hydrogenases. The [FeFe]-hydrogenase enzyme contains a di-iron active centre (HydA), known as the H-cluster, has as a [4Fe–4S] subcluster linked by a cysteine thiolate to a modified [2Fe] subcluster, known as the H-cluster. Catalysis of H_2 takes place at the di-iron centre at the [2Fe] subsite. There are three conserved cysteine-motifs which coordinate the active site H cluster, known as P1, P2 and P3 for [FeFe]-hydrogenases. These [2Fe] subsites and non-protein ligands are synthesized by the hydrogenase maturation enzymes: HydE, HydF and HydG[125]. The mechanisms of

this multienzymatic biosynthesis required for maturation is poorly understood. The [2Fe] subsite is first assembled onto a maturation enzyme, HydF. Which is then delivered to the apo-hydrogenase for activation[126]. The chemical site of [FeFe]-hydrogenases do not vary, however, the residues in the motifs binding the [FeFe] centre of diverse groups can vary[66].

1.4 Biogeography

Microbial biogeography is a relatively new field, with the bulk of foundation studies completed in the early 2000's[127]. This field follows macro-geographical guidelines, with caveats, to determine which ecological processes and factors modulate and generate microbial diversity, including speciation, dispersal, diversification, genetic drift and extinction events[127,128]. Biogeography is the distribution of biodiversity over space and time, aiming to reveal where organisms live and why[129]. Microbial ecology is a complimentary (and sometimes overlapping) field to biogeography which elucidates microorganisms' relation to each other and their physical surroundings, including characterisation of species, ecological niche, energy requirements and response to environmental perturbations[130].

Rapid improvements in cultivation-independent molecular techniques have revolutionised the fields of microbial ecology and biogeography[131]. In the last ten years, a large bulk of microbial diversity has been understood through 16S rRNA gene community profiling and more recently, metagenomics. This increase in popularity is partially due to high-profile projects such as the human microbiome project[132]. Next-generation sequencing technologies such as the Illumina platform offer high-quality sequencing and affordable rates for both 16S rRNA gene and metagenomic sequencing[133]. Simply, DNA from an environment can be sequenced and community diversity can be compared with the repertoire of public databases including Sequence Read Archive (SRA) and GenBank (NCBI) or European Nucleotide Archive (ENA - EBI)[134,135]. With additional physicochemistry and metadata (such as elevation, spatial distance), the factors influencing microbial community composition can be elucidated statistically using biogeographical analyses[52].

For dynamics and functionality of a community, "meta-omics" techniques can be employed in tandem using DNA, RNA, protein and metabolites, by "meta-omics" techniques metagenomics, metatranscriptomics, meta-proteomics and metabolomics, respectively[136]. These advances in technology, and the ensuing ease and affordability, have contributed to the shift towards studying metagenomics and metatranscriptomics of microbial communities to elucidate a snapshot of the functional capability, redundancy and actively transcribed function of any given environment. More studies also cite the need for temporal studies of ecosystems to understand the factors which modulate diversity and the robustness of community function. Genetic advancements have allowed the viewing of microbial diversity in habitats from Antarctica to the Arctic[137,138]. Studies have been accomplished for human gut microbiomes[139], the plant phyllosphere[16], acid mine drainage[140], terrestrial soils[141], cold and hot desert soil and rocks[142], Antarctic dry valleys[1,39,143], arctic soils[137], geothermal springs[52], deep sea hydrothermal vents[144], alpine zones[145], volcanic soils[146,147], volcanic rock surfaces[3,148], surface oceans[149], bathypelagic ocean zones[150] and many built environments[151]. Notably, these studies are not limited to natural ecosystems with recent studies examining constructed microcosms which were cultivated in simplified synthetic media to understand community assembly. These experiments were used to infer function and determine whether community composition is predominantly governed by nutrient availability[152].

There is a wealth of knowledge about the overlap, and in contrast, the endemic nature of microbial communities at a global scale[153,154] and the ecology of microbes for many natural and engineered environment is known.

1.5 Literature Review Summary

Extreme environments harbour a great deal more microbial diversity than previously appreciated. However, the mechanisms for survival of extremophilic microorganisms when faced with ecological conditions outside their ideal growth conditions remain elusive. Recent studies have outlined that persisting heterotrophs, Antarctic communities, Hawaiian volcanic communities and thermophilic isolates can utilize trace gases from the troposphere to persist in starvation[1,3,44,45,66]. Indeed, many microbial communities also survive in environments which lack of traditional primary producers (phototrophic microorganisms and plants) by using the mechanism of trace gas oxidation[1]. As volcanic ecosystems away from active vents are often devoid of overt carbon sources and/or phototrophic communities (micro- or macro-) an obvious next question is: ***could the reduced flux or passive degassing of H_2 and CO from volcanoes support microbial communities employing high affinity trace-gas hydrogenases and carbon monoxide dehydrogenases as a means for persistence?***

To answer this question, I will undertake a molecular and microcosm-based research project on volcanic soils away from active vents sites from various locations in the TVZ.

1.6. Hypothesis

The microbial ecology of representative New Zealand non-active volcanic soils will be similar to other oligotrophic soils systems. In bare, volcanic and alpine soils, trace-gas (H_2 and CO) oxidation is used as a mechanism of persistence in nutrient limitation and is a valuable source of supplemental energy for microbial taxa.

1.7. Statement of Research Questions and Objectives

1. **To describe the microbial ecology of Tongariro National Park, White Island, Mt. Urchin and Mt. Sugarloaf**

   Objective: Use bioinformatics to determine microbial diversity and abundance. This can be done by producing operational taxonomic units (OTUs) to show the relative abundance of microbial taxa in soil from South, Red and Central Craters of Mt. Tongariro at 0 cm and 10 cm. Additional volcanic and non-volcanic bare soil sites from White Island, Mt. Urchin Track and Mt. Sugarloaf at depths of 0 cm and 10 cm will be used as control sites.

2. **To outline the oxidation of atmospheric trace gases (H_2, CO) as a mechanism of microbial persistence**

   Objective: To use gas chromatography to track oxidation of trace gas of CO and H_2 in volcanic soil samples, using biological and technical triplicates at 4 °C. Samples from volcanic soils of White Island (Whakaari), Mt. Tongariro – South, Red and Central Craters, alpine and bare soils from Mt. Urchin (a non-volcanic mountain adjacent to the Taupō Volcanic Zone (TVZ)) and soils from Mt. Sugarloaf from New Zealand's South Island were selected as controls. To screen soils for high-affinity enzymes responsible for the oxidation of H_2 and CO at atmospherically relevant concentrations.

3. **To demonstrate persistence of thermophilic taxa in soils of Tongariro National Park**

   Objective: Demonstrate trace gas oxidation occurs at thermophilic temperatures by incubating soil and thermophilic isolates at 60 °C and analysing gas composition by gas chromatography.

Chapter 2: Microbial Ecology and Biogeography of Mt. Tongariro, White Island, Mt. Urchin and Mt. Sugarloaf

2.1 Introduction

The aim of Chapter 2 is to explore the diversity of the microbiota of Mt. Tongariro soils and to compare these results to other similar volcanic, high elevation, and oligotrophic systems. Controls from non-volcanic bare soils from Mt. Urchin and Mt. Sugarloaf are also explored. White Island is used as an additional volcanic soil control. The volcanic areas (Mt. Tongariro and White Island) of this study are andesitic stratovolcanoes (Figure 2)[5]. Mt. Urchin, the non-volcanic mountain North Island control is located in the Kaimanawa Ranges on the Eastern boundary of the TVZ (approximately 21 km from Mt Tongariro; Figure 3) and is composed predominantly of greywacke and formed as a result of tectonic rather than volcanic activity. The South Island control, Mt. Sugarloaf is adjacent to the Southern Alps and predominantly consists of sandstone and argillate (sedimentary rock types)[155]. This study represents the first investigation into the microbial diversity of Tongariro National Park. There have been a limited number of previous biological studies in this region which investigated the presence of vascular plants within the alpine zone[156] or have characterised microorganisms isolated from volcanic soils of Mt. Ngauruhoe[50,157]. Previous studies of bare, volcanic soil in non-TVZ locations have shown that these often oligotrophic, extreme volcanic and environments harbour a surprising amount of microbial abundance and diversity[1,158]. One study of note reported shifts in microbial diversity towards communities dominated by taxa from the phylum Actinobacteria in Atacama high alpine volcanic soils which have extremely low in moisture content[37].

To test the differences in community composition between soils at the surface and soils at depth, soil samples were collected at two depths (0 cm and 10 cm): total environmental DNA was extracted, sequenced and processed using a custom bioinformatics pipeline for the 1000 Springs Project[52]. To assess the influence of the physical, geographical and the geochemical conditions on microbial community composition, the physicochemistry of the soil matrices were analysed for 25

chemical species, pH and temperature, along with geographical metadata, and tested against alpha and beta diversity metrics.

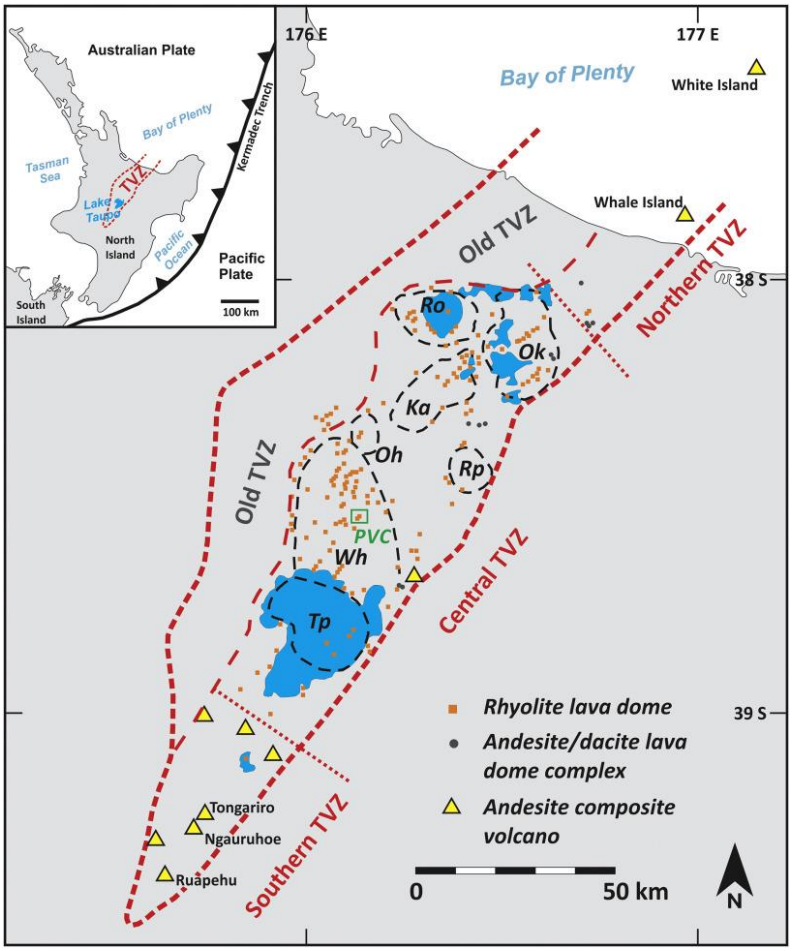

Figure 2: Map of the Taupō Volcanic Zone (TVZ), New Zealand showing the location of Mt Tongariro and White Island (Whakaari). Map sourced from Kosik et al., 2017. [5].

Mt. Urchin, the non-volcanic North Island control is located in the Kaimanawa Ranges on the Eastern boundary of the TVZ (Figure 3) and is composed predominantly of greywacke. The South Island control, Mt. Sugarloaf is adjacent to the Southern Alps and predominantly consists of sandstone and argillate (sedimentary rock types)[155].

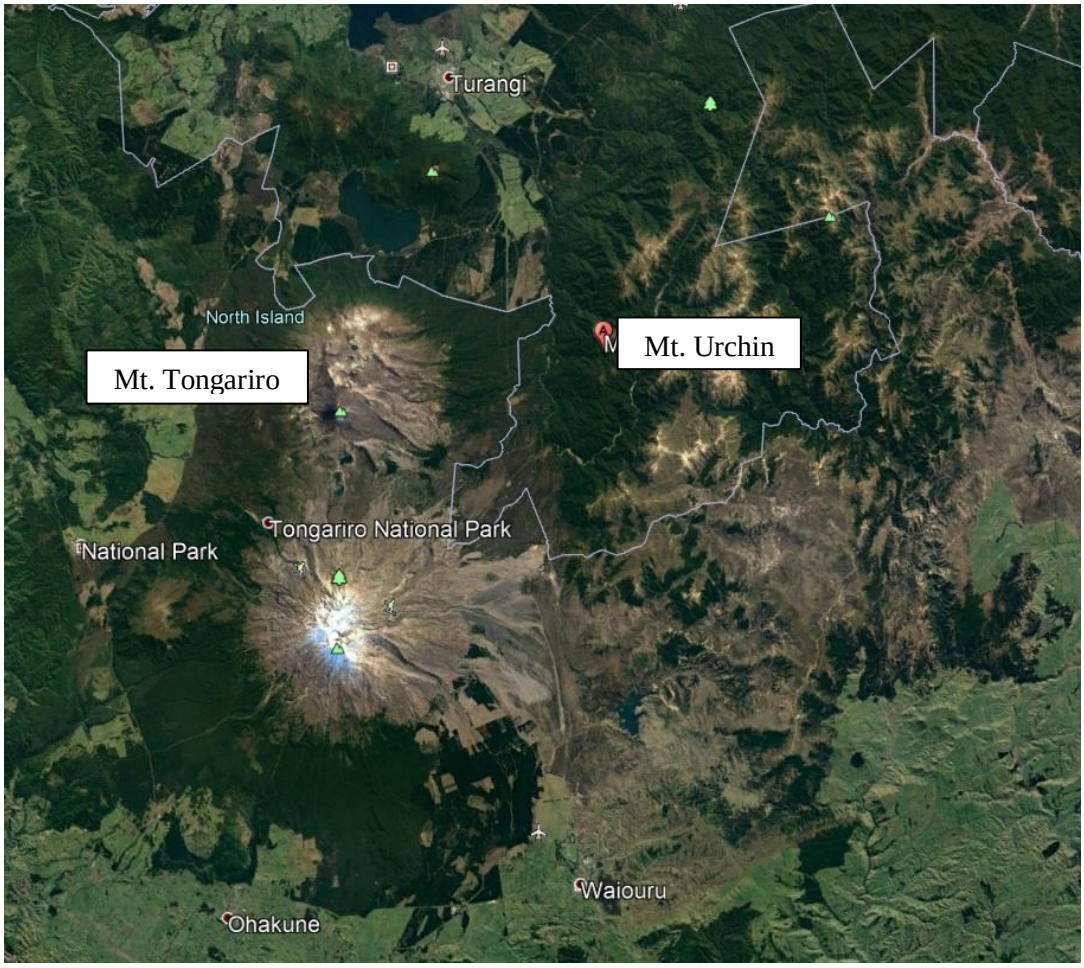

Figure 3: Satellite image of Tongariro National Park showing the geographical location of Mt Urchin, Kaimanawa Forest Park. The South, Central and Red Crater sample sites are located on the Mt Tongariro Massif (see Figure 1 for detail). Image from Google Earth (Pro v7.1, 18/02/2015, 30°10'38.17" S 175°51'05.68" E), viewed 22/01/2019.

2.2 Materials and Methods

2.2.1 Field Sampling and processing

In total, soil samples (n=30) were collected between May 2017 and March 2018 from four sites in the TVZ: South, Red and Central Craters of Mt. Tongariro; one non-volcanic mountain sample from Mt. Urchin in the Kaimanawa Forest (north-east of the TVZ boundary) and one additional non-volcanic mountain site from the South Island (Mt. Sugarloaf). All samples were collected from areas of bare soils above the tree line (in the Alpine Zone) devoid of root systems and vascular plants. Gas samples were taken at the surface (0 cm) and at depth (10, 30, or 50 cm) using a custom-made gas sampler (GNS Science, Wairakei Research

Centre, New Zealand). The soil gases were collected and stored in 9 ml Tedlar gas sampling bags (Tedlar®, Sigma-Aldrich, USA) for analysis by gas chromatography (GC). Soil samples were collected using an ethanol-sterilised trowel and/or spatula. Technical triplicates for the control sites and biological triplicates for South, Red and Central Craters of Mt. Tongariro. Triplicates of ~ 20 g of soil was transferred into 50 ml falcon tubes using a trowel which was sterilised using 70% w/v ethanol spray. Samples were taken from surface (0 cm) and depth (10 cm), and were frozen at -80 °C immediately upon arrival at the laboratory. Total environmental DNA was extracted using the NucleoSpin® soil kit (Macherey-Nagel, Düren, Germany) using the manufacturer's specifications for extraction method, with 125 µl of 7.5 M potassium acetate added as an additional inhibitor removal step for the standard protocol. Soil DNA was extracted in triplicate using an initial 0.25-0.3 g soil weight and then triplicates were pooled together and stored at -20 °C.

2.2.2 Physicochemical Analysis

Twenty-five physicochemical parameters were determined for each soil sample. Inductively coupled plasma–mass spectrometry (ICP-MS) was used to determine the concentrations of metals and non-metals (eight species; Supplementary Table 2) and inductively coupled plasma-optical emission spectrometry (ICP-OES) was used to determine the concentrations of non-metal and metals for seven species. Ion chromatography methods were used to determine chloride (Cl^-), nitrate (NO_3^-) N and SO_4^{2-}. FIA was used to determine NH_4^+ (as N) and dissolved reactive phosphorus (New Zealand Geothermal Analytical Laboratory (NZGAL), Wairakei Research Centre, New Zealand). Total organic carbon (TOC) and total inorganic carbon (TIC) were determined by catalytic combustion at 900 °C (Hills Laboratory, Hamilton, New Zealand). Soil moisture content was calculated using gravimetric method in which 1 g of soil is weighed, then dried to constant weight at 90 °C[159,160]. Soil pH was measured by dissolving 1 g of soil in 10 ml ddH_2O using a Hanna Instruments Edge pH meter to measure pH (Woonsocket, RI, USA), temperature (TEMP) recorded using a Fluke 51-II thermocouple (Fluke, Everett, WA, USA). Physicochemical methods can be found in Supplementary Table 1.

2.2.3 PCR and Sequencing

The V4 region of the 16S rRNA gene was amplified in triplicate reactions which were the pooled using universal Earth Microbiome Project primers F515 (5'-GTGYCAGCMGCCGCGGTAA-3'), R806 (5'-GGACTCANVGGGTWTCTAAT-3')[48]. PCR master mixture per reaction contained: 28.5 µl of PCR H_2O, 7 µl of PCR enhancer, 5 µl dNTP, 5 µl 10X Taq buffer, 0.5 µl of each primer, 0.5 µl of Taq (5U μl^{-1}) and 3 µl of template DNA (50 ng DNA) using the i-Taq™ DNA Polymerase kit (iNtRON Biotechnology, South Korea).

PCR protocol: Initial denaturation was achieved at 94 °C for 5 min, followed by 36 PCR cycles consisting of denaturing for 45 s at 94 °C, annealing for 30 s at 55 °C and extension for 60 s at 72 °C. A final extension was performed for 5 min at 72 °C, using a T100™ Thermal Cycler (Bio-Rad, Hercules, CA, USA). To confirm the identity of PCR products, bands were visualised by gel electrophoresis, using 1.5% v/v SeaKem® LE Agarose (Lonza, Rockland, ME, USA). Bands of the correct product size (~250 bp) were aseptically cut from the gel and purified using NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel, Düren, Germany). Amplicon sequencing of the V4 region was performed using the Illumina MiSeq platform (Macrogen, Geumcheon-gu, Seoul) using Agilent Herculase II Fusion DNA Polymerase (Agilent, Santa Clara, CA, USA) for PCR reagent and Illumina MiSeq Reagent Kit v2 for chemistry (Illumina, University City, San Diego).

2.2.4 DNA sequence processing

DNA sequences were processed using a custom 1000 Springs Project pipeline[52] which utilises functions in UPARSE[135], QIIME[161] and mothur[162]. Sequences were demultiplexed and then analysed for quality control and adapter sequences were removed using trim_galore, a wrapper tool for cutadapt and fastqc (https://github.com/FelixKrueger/TrimGalore)[163]. Files were then converted using "from fastq to fasta format" using USEARCH v7 (Part of the UPARSE package), with maximum error of 1.0% (fastq_maxee 1.0) to filter out low quality reads[135]. End-paired reads were merged to give more confidence to the Q score of each base call in UPARSE. All samples were concatenated in series. Using mothur, short (< 200 bp) and long (> 500 bp) sequences were trimmed and homopolymers (> 6) were

removed[162]. The total final read count was 6,043,458 with a mean of 188,858 reads per samples. A *de novo* database was then constructed using USEARCH v7. All replicates and singletons (based on an abundance of < 2 reads) were removed during processing. Operational taxonomic unit (OTU) sequences were clustered at 97% similarity and assigned to taxonomy using the Ribosomal Database Project Classifier (RDP)[164], with a minimum confidence score of 0.5 (50%) against the Silva 16S v132 database[165]. Taxa were then filtered to remove unwanted chloroplasts and mitochondria sequences in QIIME and rarefied to the lowest sequence count detected in the sample summary (73,000 sequences). This resulted in the removal of two blank samples (total $n = 30$).

2.2.5 Statistical Analyses Methods

All statistical analyses and visualisations were performed in R (v3.5.2) using Phyloseq (v1.26.1)[166], Vegan (v2.5.3.)[167] and ggplot2 (v3.1.0) packages[168]. Taxonomy was assigned to OTUs using the RDP classifier against the SILVA 16S rRNA database (v132) with a minimum confidence score of 0.5 (50%). Using phyloseq, OTUs were agglomerated to the corresponding phylum using the tax_glom function in phyloseq, with phyla containing <0.01% read abundance removed. Relative abundance at the phylum level for each site was then presented in stacked bar graphs.

Alpha diversity metrics were calculated using the estimate_richness function in the phyloseq package and graphed using ggplot2. First, a best fit linear regression model was used to identify the relationship between Shannon distance indices (H) and each individual physicochemical parameter. This model was also repeated for OTU richness (Observed OTUs) against pH and soil temperature only, to confirm the lack of relationship between alpha diversity and these two variables.

A multiple linear regression (MLR) model was then calculated using the significant values ($R^2 = <0.05$) from single linear regression against the Shannon diversity index. Akaike information criterion (AIC) was subsequently applied to filter the variables and improve the fit of the MLR model[169]. Non-parametric

Kruskal-Wallis (*H*) testing was used on categorical variables 'site' and 'depth' against Shannon diversity to identify any significant influences on diversity. Finally, correlation between significant physicochemical parameters and Shannon diversity index (H) was performed with Pearson's coefficient |*r*| using the cor.test function in R[170].

Beta diversity was analysed using Bray-Curtis dissimilarity and displayed using non-metric multidimensional scaling (NMDS). Mantle tests, using Spearman's correlation coefficient (ρ) with 999 permutations were run for each physicochemical parameter against beta diversity. In decreasing order of correlation, metadata were added to permutational multivariate analysis of variance (PERMANOVA) analysis using the adonis function in vegan. Constrained correspondence analysis (CCA) was performed for beta diversity (again using Bray-Curtis dissimilarity) against physicochemistry (using the cca function in vegan) which was plotted using ggplot2. Continuous physicochemistry were added to the model in order of highest to lowest Mantel result. For the phylogenetic tree, representative OTUs were aligned using PyNAST[171] and an approximately maximum-likelihood phylogenetic tree was generated with FastTree[172]. OTUs were then agglomerated to their respective genera (using the tax_glom function in phyloseq), and low abundant taxa (<1% average relative abundance) were removed for visualisation. Where taxonomy was unable to be assigned to genera, the closest taxonomic rank is shown (p = phylum, c = class, o = order, f = family)[173].

2.3 Results: Microbial Ecology and Biogeography

2.3.1 Site Description

This study focused on four sites within the TVZ, New Zealand: South Crater, Red Crater and Central Crater of Mt. Tongariro, one site on White Island (Whakaari) and two additional non-volcanic controls from Mt. Urchin (Kaimanawa Ranges, adjacent to the TVZ) and Mt. Sugarloaf, South Island. At each site, triplicate soil samples were collected at 0 and 10 cm depth. Geographic coordinates for sites are defined in WGS (World Geodetic System, 1984) south-east decimal format. Sample site descriptions and metadata are presented in Table 1 and images of the sample sites are presented in Figures 4-9.

Chemical methods used to determine physiochemical species can be found in Supplementary Table 1. Physicochemical metadata can be found in Supplementary Table 2

location	date	sample number	pH	temp	altitude (m)	geographical coordinates	light (lux)	air temp	air pressure (hPa)	general description
Mt. Urchin Kaimanawa	7/05/2017	EQC_1_0cm EQC_1_10cm	5.4 5.3	7.7 4.9	1370	(-39.17549501 175.85177805)	12032	7.7	991.80	Area zone clear of plants (1m). Accumulation of soil on small plateau. Moss, grasses and small shrubs only. No wind, clear day. Difficult to find a clear spot where we could sample at depth. Shale-like rock everywhere.
South Crater Mt. Tongariro	15/05/2017	EQC_2_0cm EQC_2_10cm EQC_3_0cm EQC_3_10cm EQC_4_0cm EQC_4_10cm	5.0 4.8 4.9 4.8 5.1 4.9	0.0 2.9 0.0 2.9 0.0 2.9	1638	(-39.14059230, 175.63978407)	11335	1.0 1.0 1.0 1.0 1.0 1.0	991.80	Fluffy soil - suspect highly aerated. Lots of ice needles up to 5 cm pushing out of the ground. Soil particles on top of needles. Devoid of plants. Not moss or lichen visible. Area looks like an alluvial plain. In cloud. Moderate wind.
Red Crater Mt. Tongariro		EQC_5_0cm EQC_5_10cm EQC_6_0cm EQC_6_10cm EQC_7_0cm EQC_7_10cm	6.1 6.3 6.4 6.4 7.0 6.4	22.8 44.3 29.6 31.4 23.0 45.6	1822	(-39.13537095, 175.65071738)	NA	0.1 0.1 0.1 0.1 0.1 0.1	991.80	Layered soil in cutting directly on Tongariro track. Dark with some red. Relatively fine grain. Clearly laid down during eruptions. The soil temperature was highly variable reaching up to 60C is some spots. V. strong winds although protected from the wind by cutting (~2m vertical height). Samples taken on a 20-30 degree slope..
Central Crater Mt. Tongariro		EQC_8_0cm EQC_8_10cm EQC_9_0cm EQC_9_10cm EQC_10_0cm EQC_10_10cm	5.4 5.5 5.4 5.4 5.3 5.5	0.3 3.0 0.3 3.0 0.3 3.0	1681	(-39.12797297, 175.65355132)	NA	1.1 1.1 1.1 1.1 1.1 1.1	991.80	Alluvial plain similar to South Crater. Some small shrubs nearby (20m). Seedlings, moss and lichen also present. Gravel a little larger than other crater sites. Moderate to strong winds. Cloud clearing.
White Island		EQC_13_0cm EQC_13_10cm EQC_14_0cm EQC_14_10cm EQC_15_0cm EQC_15_10cm	3.2 3.9 2.9 4.1 2.9 3.8	23.3 26.7 22.5 27.4 21.3 24.3	17	(-37.52316091, 177.18765109)	24060	17.4 17.4 17.4 17.4 17.4 17.4	1064.80	North of Sulphur Cave, that's north west of Fumarole Zone. About 30m from tourist path. Next to Geonet gas peg #68

Mt. Sugarloaf South Island	26/03/2018	CAS_1_0cm	7.5	17.8	1160	(-43.03272223, 171.78344444)	NA	16.4	NA	Sample site on open patch near scree slopes. Some vegetation nearby. A lot of roots @ -1-5cm. Soil damp/moist but not water logged.
		CAS_1_10cm	7.0	14.9				16.4		
		CAS_2_0cm	6.6	17.2	1070	(-43.03272223, 171.78344444)		16.4		Sample site on scree slope (+30 Degree angle). 25 m from closest plant. No roots observed. Soil moist/damp, but not waterlogged.
		CAS_2_10cm	6.4	15.2				16.4		

Table 1: Sampling metadata including location, date, sample identifier, pH, temperature (°C), altitude (m), geographical location (World Geodetic System 1984), light (lux), air pressure (hPa) and general site description. NA = not applicable.

Figure 4: Photographs showing sample site and gas apparatus, White Island (Whakaari), TVZ.

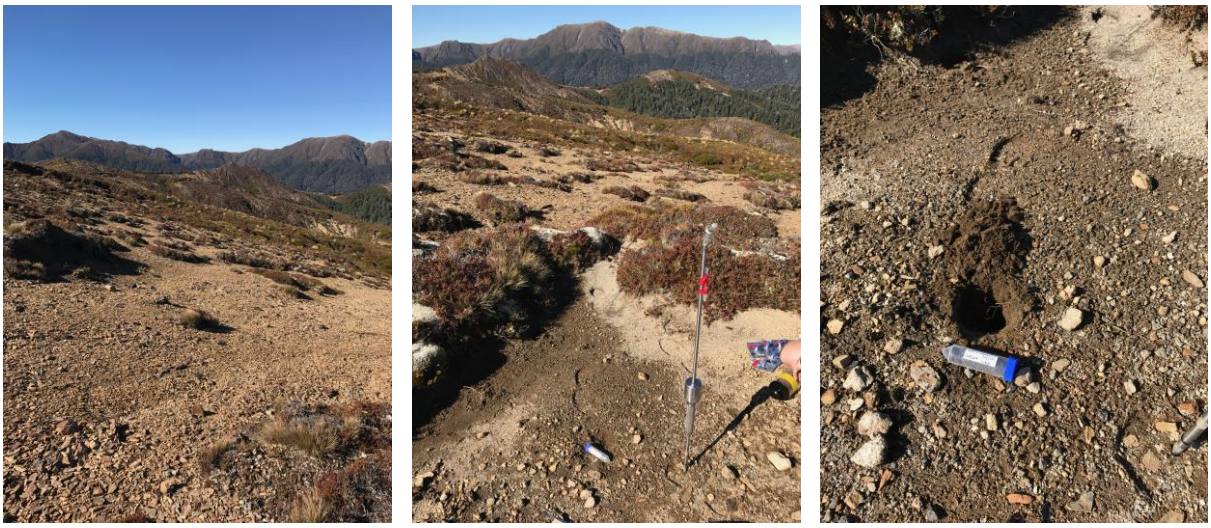

Figure 5: Photographs showing sample sites and gas apparatus, Mt Urchin, Kaimanawa Forest Park.

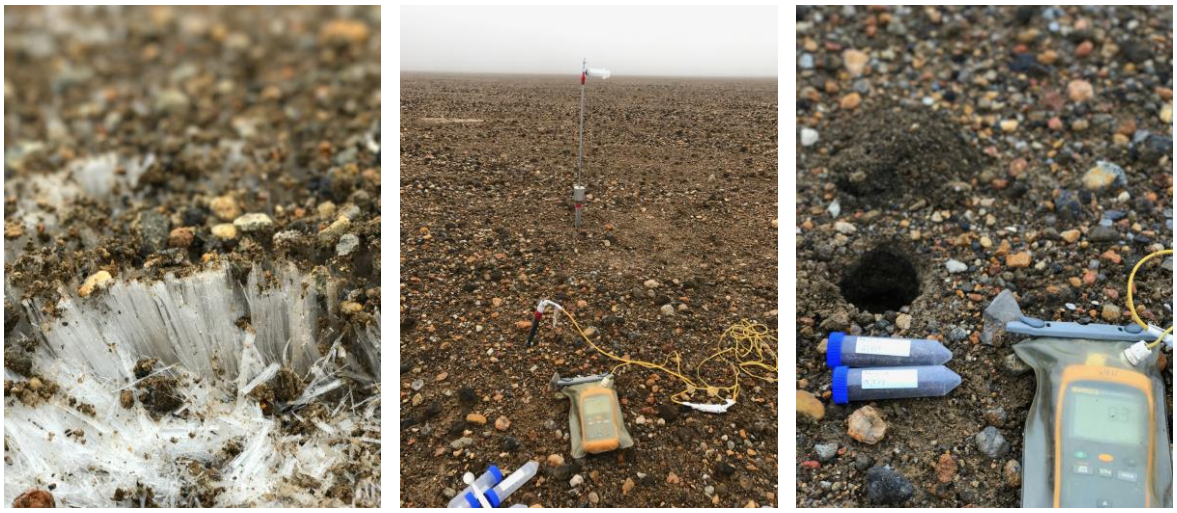

Figure 6: Photographs showing sample sites and field and gas apparatus, South Crater, Mt. Tongariro, TVZ.

Figure 7: Photographs showing sample sites and field and gas apparatus, Red Crater, Mt Tongariro, TVZ.

Figure 8: Photographs showing sample sites and field and gas apparatus, Central Crater, Mt. Tongariro, TVZ.

Figure 9: Photographs showing sample sites and field and gas apparatus of Mt Sugarloaf, CASS Station, South Island. Adjacent to the Arthur's Pass, Southern Alps, New Zealand.

2.3.2. Results: Microbial Community Composition

To understand the soil community composition of Mt. Tongariro, Mt. Urchin, White Island and Mt. Sugarloaf, the total DNA content from the sampled soils was extracted and sequenced using EMP primers (see Section 2.2.4). Downstream bioinformatical analysis of the 16S rRNA gene was performed, in order to produce an OTU table. After OTU sequences were assigned to taxonomy, the mean relative abundance (>0.01%) was calculated to show diversity. Original OTU sequences were also used for alpha and beta diversity measures.

Below, Figure 10 shows the most abundant phyla (n=17, > 0.1% average relative abundance) in the entire OTU dataset: Acidobacteria, Actinobacteria, Armatimonadetes, Bacteroidetes, Chloroflexi, Crenarchaeota, Cyanobacteria, Deinococcus-Thermus, Elusimicrobia, Epsilonbacteraeota, candidate phylum GAL 15, Gemmatimonadetes, Planctomycetes, Proteobacteria, Thaumarchaeota, Verrucomicrobia and candidate phylum WPS-2. The relative levels of these phyla in each of the site triplicates for South, Red and Central Crater are even. The difference in composition is more noticeable between sites than at sites. Mt. Sugarloaf has much more Verrucomicrobia than the other sites and less Proteobacteria. Central Crater, both at surface (0 cm) and depth (10 cm), has more Chloroflexi, a known thermophilic phylum, than the other volcanic sites (South Crater, Red Crater, White Island). White Island and Red Crater (10 cm) lack WPS-2 (Figure 10).

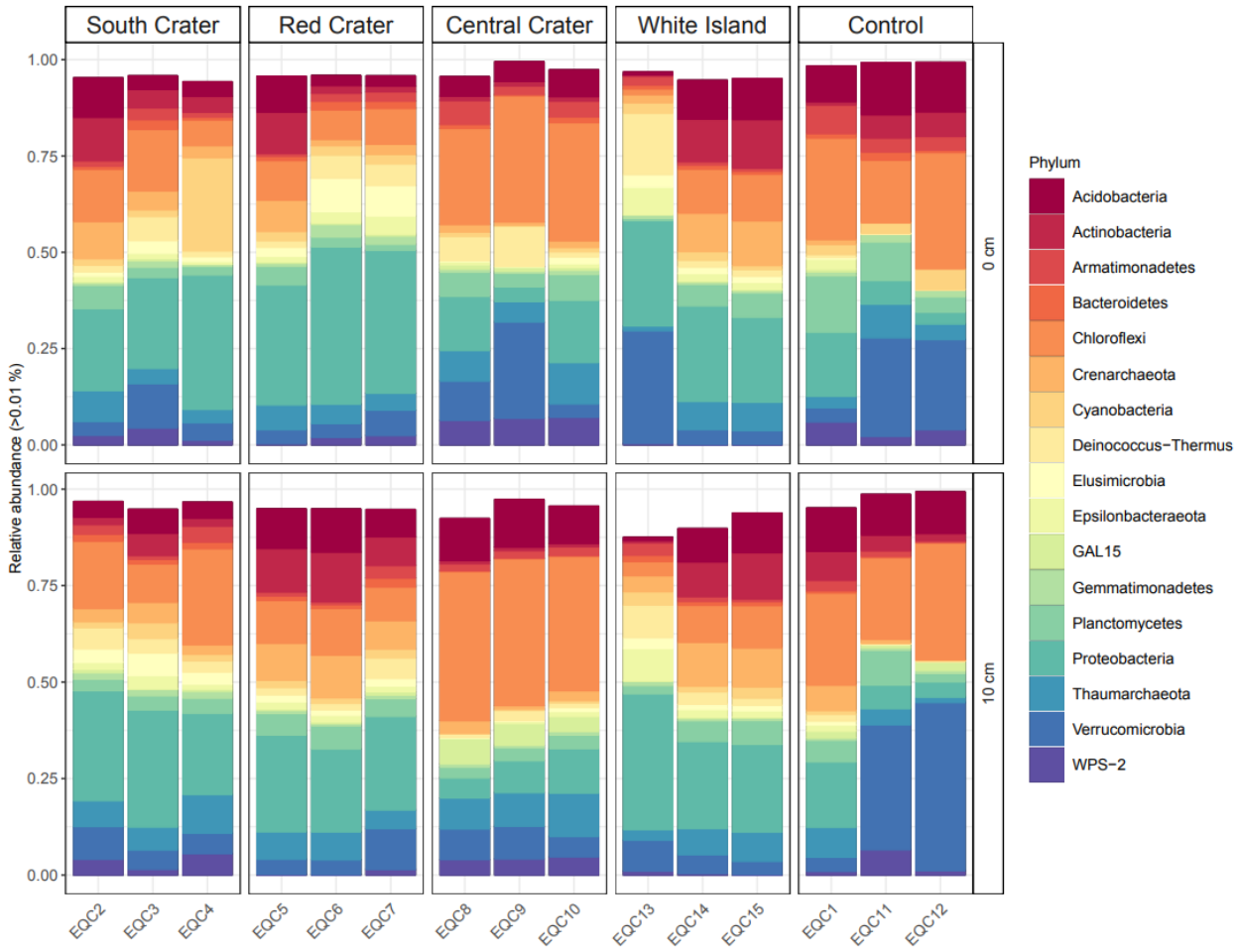

Figure 10: Bar graph of abundant phyla, displayed per sample site and depth. Only phyla with >0.01% average relative abundance are shown (n=17, coloured). The individual community compositions of each sample are split by location (South Crater, Red Crater, Central Crater of Mt. Tongariro, TVZ, shown as biological triplicates), biological triplicates of White Island soil (Whakaari), and non-volcanic control samples from Mt. Urchin (EQC1) and Mt. Sugarloaf (EQC11, EQC12), and depth (0 and 10 cm). Relative abundance is based on bacterial and archaeal OTUs of the V4 region from 16S rRNA gene sequencing, agglomerated to their respective phylum.

The phylogenetic relationship between the most abundant taxa (mean abundance >1%) across the whole dataset is presented in Figure 11 and Supplementary Table 3. These data show that all 27 abundant taxa are found in nearly all sample sites, but in varying abundance.

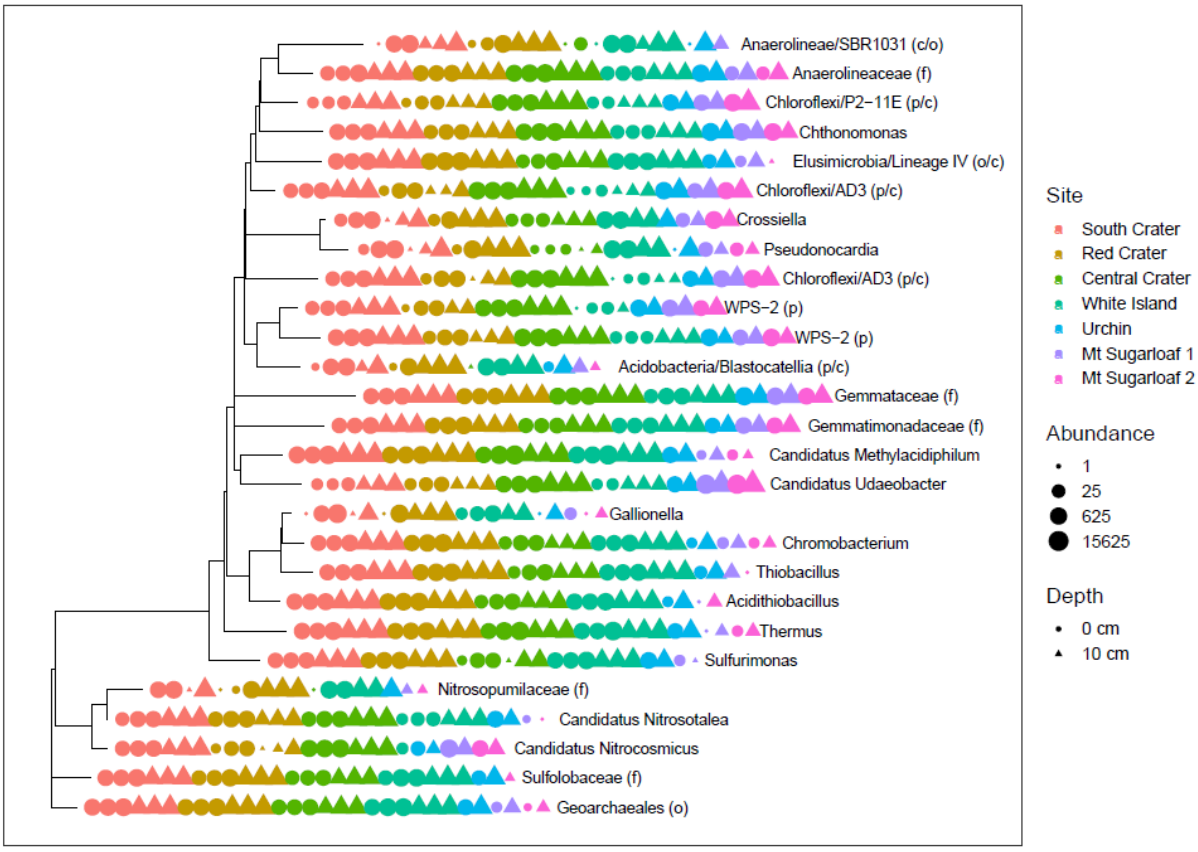

Figure 11: Phylogenetic tree depicting abundant and prevalent taxa. Representative OTUs were aligned using PyNAST and an approximately maximum-likelihood phylogenetic tree was generated with FastTree. OTUs were then agglomerated to their respective genera, and low abundant taxa (<1% average relative abundance) were removed for visualisation. Where taxonomy was unable to be assigned to genera, the closest taxonomic rank is shown (p = phylum, c = class, o = order, f = family, no rank = genus). Relative read abundance (log scale) and prevalence (presence or absence at each site) of each genus are shown in size and colour respectively, with depth displayed by shape (0 and 10 cm).

To examine the diversity between samples from the same site, alpha diversity metrics (Shannon diversity index and Observed OTU richness) were calculated using the original OTUs. No filtering was applied to the sequence data before alpha diversity indices were measured. As seen in Figure 12, the main trend for OTU richness between depths is that richness is lower at 10 cm depth for Central Crater, White Island and Mt Sugarloaf. Shannon diversity index (H) values are generally higher at 10 cm depth, meaning increased diversity and evenness exist at this depth. Values for 0 cm are more variable between biological triplicates for each volcanic site (South, Red, Central Crater, Mt. Tongariro and White Island).

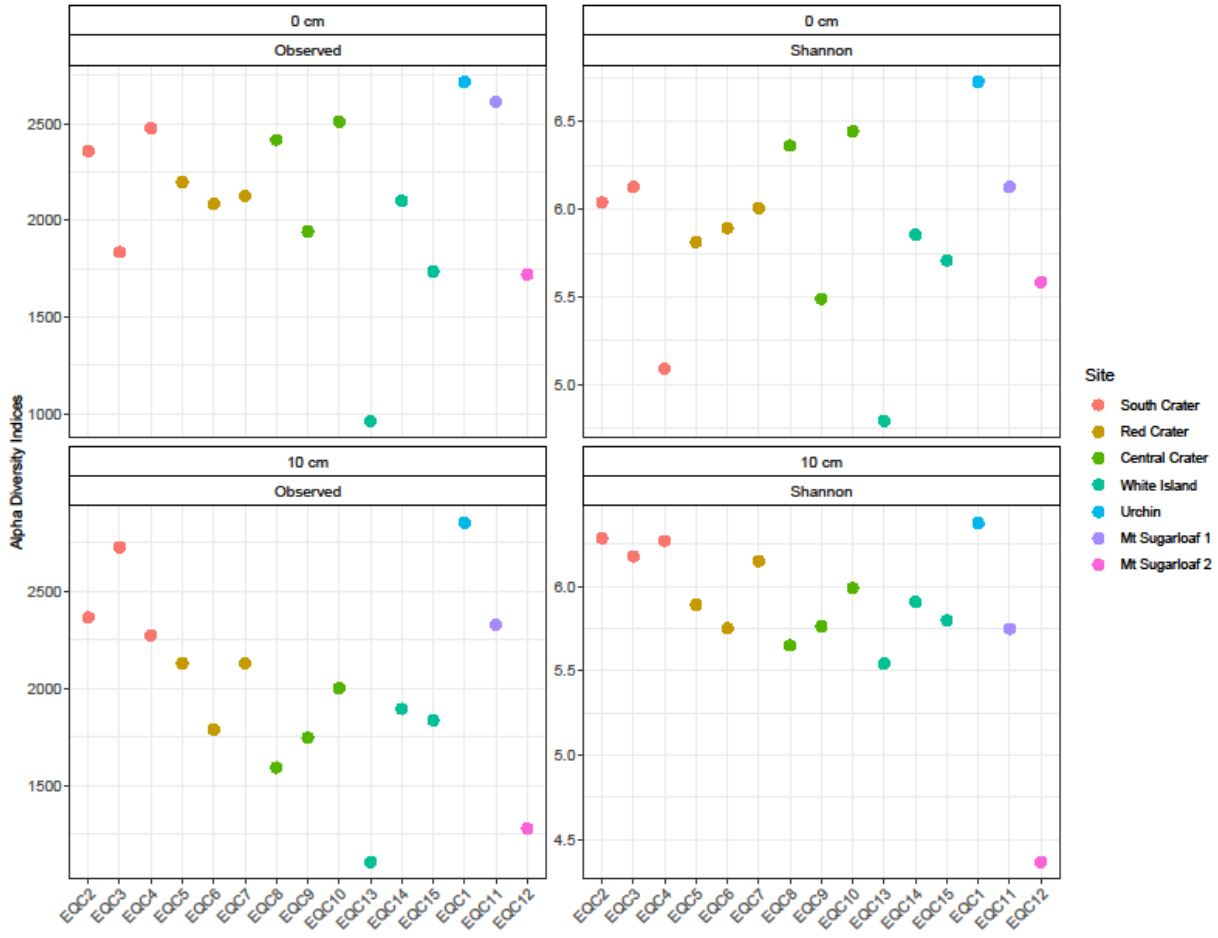

Figure 12: Alpha diversity measures of samples. Diversity is shown by both OTU richness (Observed OTUs) and evenness (Shannon diversity index, H) for each site. Samples are coloured according to location and are split by depth (0 and 10 cm). Mt. Sugarloaf has two recorded samples, compared to all other sites which have one.

Next, the alpha diversity indices (Shannon and Observed) were calculated by depth to see if depth had any effect on observed and Shannon diversity (Figure 13). When samples are split by depth, observed OTU richness at the surface (0 cm) is greater on average (shown by the mean bar in the boxplot) than at depth (10 cm). The spread of values shows that there is a larger range of OTU richness at 10 cm depth. Shannon diversity index (H) show that both at surface (0 cm) and depth (10 cm), the mean H values are similar, yet there is greater spread of values at 0 cm. This indicates that some samples (shown as outliers) have low diversity, and community structure is possibly skewed by a dominant species i.e. the abundance of taxa in the community are uneven (lower H index). Conversely, higher H index samples show that the community is both more evenly distributed and diverse. Overall, alpha diversity is higher in samples at the surface (Figure 13).

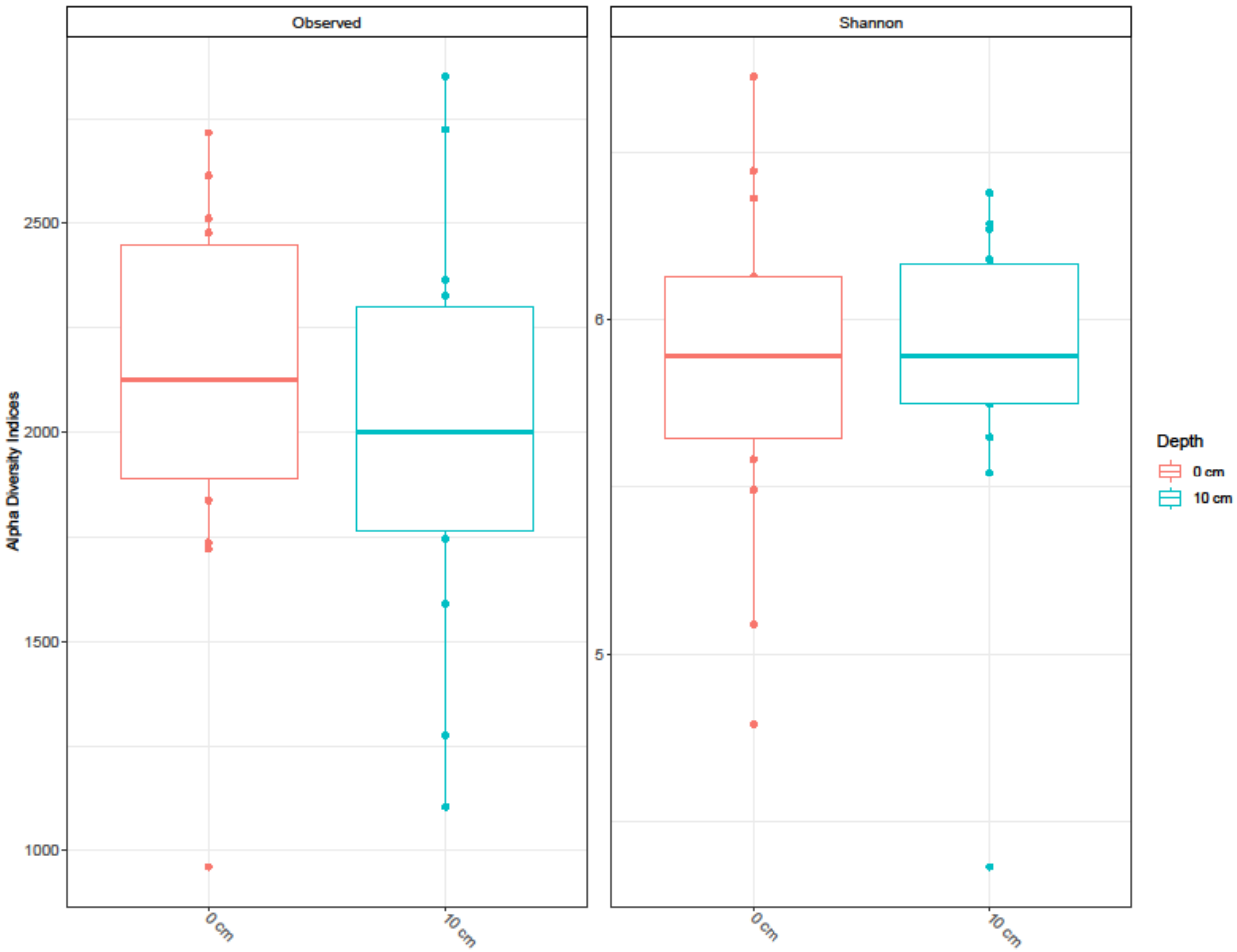

Figure 13: Boxplots visualising alpha diversity spread by soil depth. OTU richness (Observed OTUs) and evenness (Shannon diversity index, H) of each sample are grouped and coloured by depth (0 and 10 cm). Boxplots show the maximum, minimum and mean values for both depths.

If alpha diversity indices are calculated for per site, irrespective of depth (Figure 14), Mt. Urchin displays the highest mean observed OTU richness overall, followed by Mt. Sugarloaf 1, South Crater, Red Crater, Central Crater, White Island and Mt Sugarloaf 2. The sites with the greatest spread are Central Crater and White Island, suggesting that these two sites have differences in observed richness within replicate samples. Mt. Urchin also displayed the greatest Shannon diversity index mean (H), followed by South Crater, Red Crater, Mt. Sugarloaf 1, Central Crater, White Island and Mt. Sugarloaf 2 (Figure 14). The spread of values for Mt. Sugarloaf 1 is larger than the other sites, indicating differences in diversity and evenness of taxa in that site. Similarly, Central Crater has a large spread indicating differences in observed diversity and evenness in soil communities.

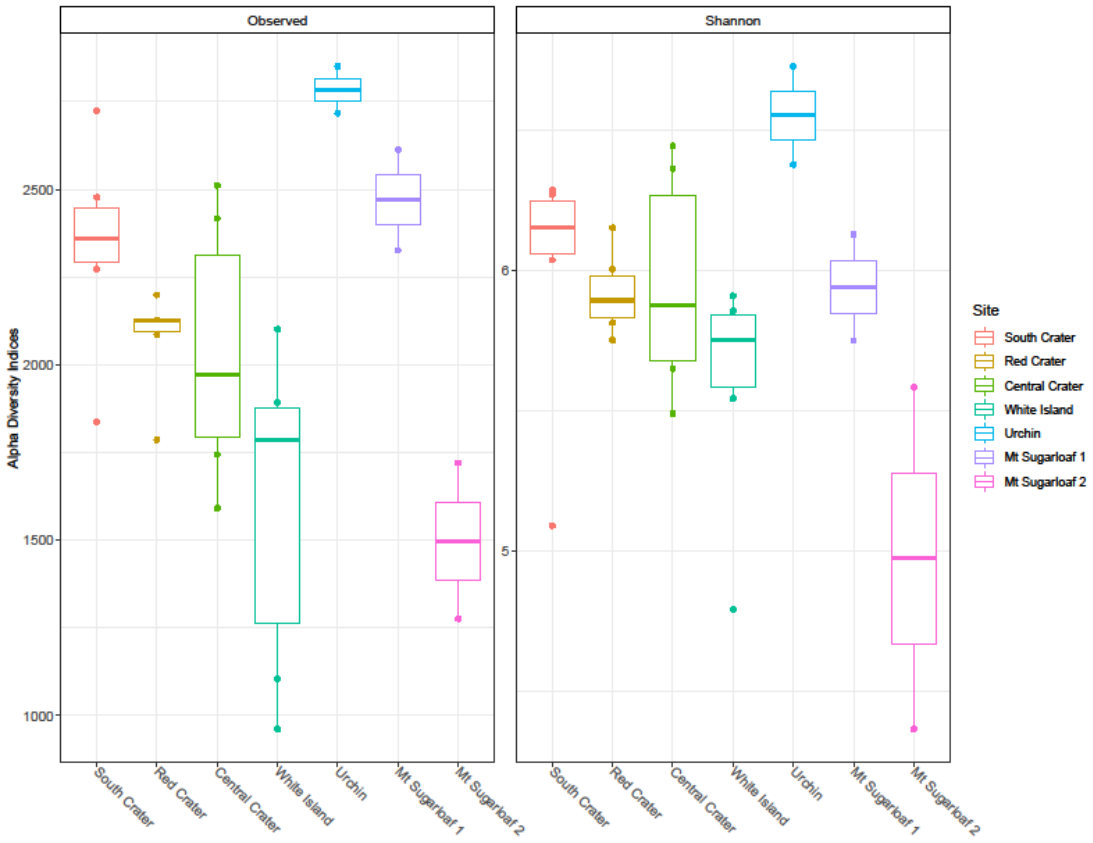

Figure 14: Comparison of alpha diversity indices based on sample site. OTU richness (Observed OTUs) and evenness (Shannon diversity index, H) of each sample are grouped and coloured by site (South Crater, Red Crater, Central Crater, White Island, Mt. Urchin, Mt. Sugarloaf 1 and Mt. Sugarloaf 2). Boxplots show the maximum, minimum and mean values for each site.

Next, to see whether soil temperature, or pH had any significant effect on alpha diversity, the linear regression model was used against alpha diversity indices (Shannon and Observed). Figure 15 shows that soil temperature had no significant effect on observed OTU richness, or Shannon diversity, according to the linear regression model (Observed: $R^2 = 0.11$, $p = 0.08$; Shannon: $R^2 = 0.03$, $p = 0.35$). Similarly, pH has no significance on observed Shannon diversity according to the linear regression model (Observed: $R^2 = 0.08$, $p = 0.12$; Shannon: $R^2 = 0.01$, $p = 0.61$) (Figure 16).

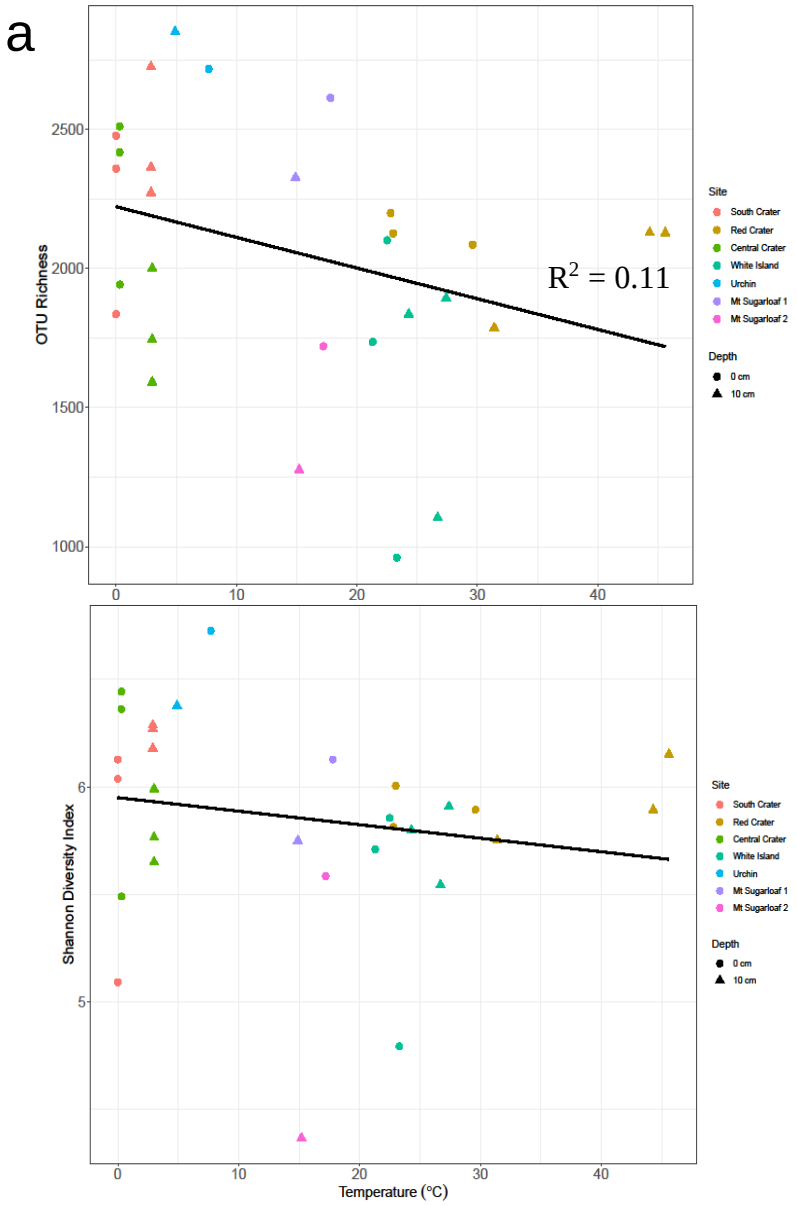

Figure 15: Comparison of temperature against alpha diversity indices. a, OTU richness (Observed OTUs) and **b**, evenness (Shannon diversity index, H) of samples are plotted against soil temperature (°C). Samples are coloured by site and shaped by depth. Linear regression was applied to these relationships (Observed: $R^2 = 0.11$, $p = 0.08$; Shannon: $R^2 = 0.03$, $p = 0.35$).

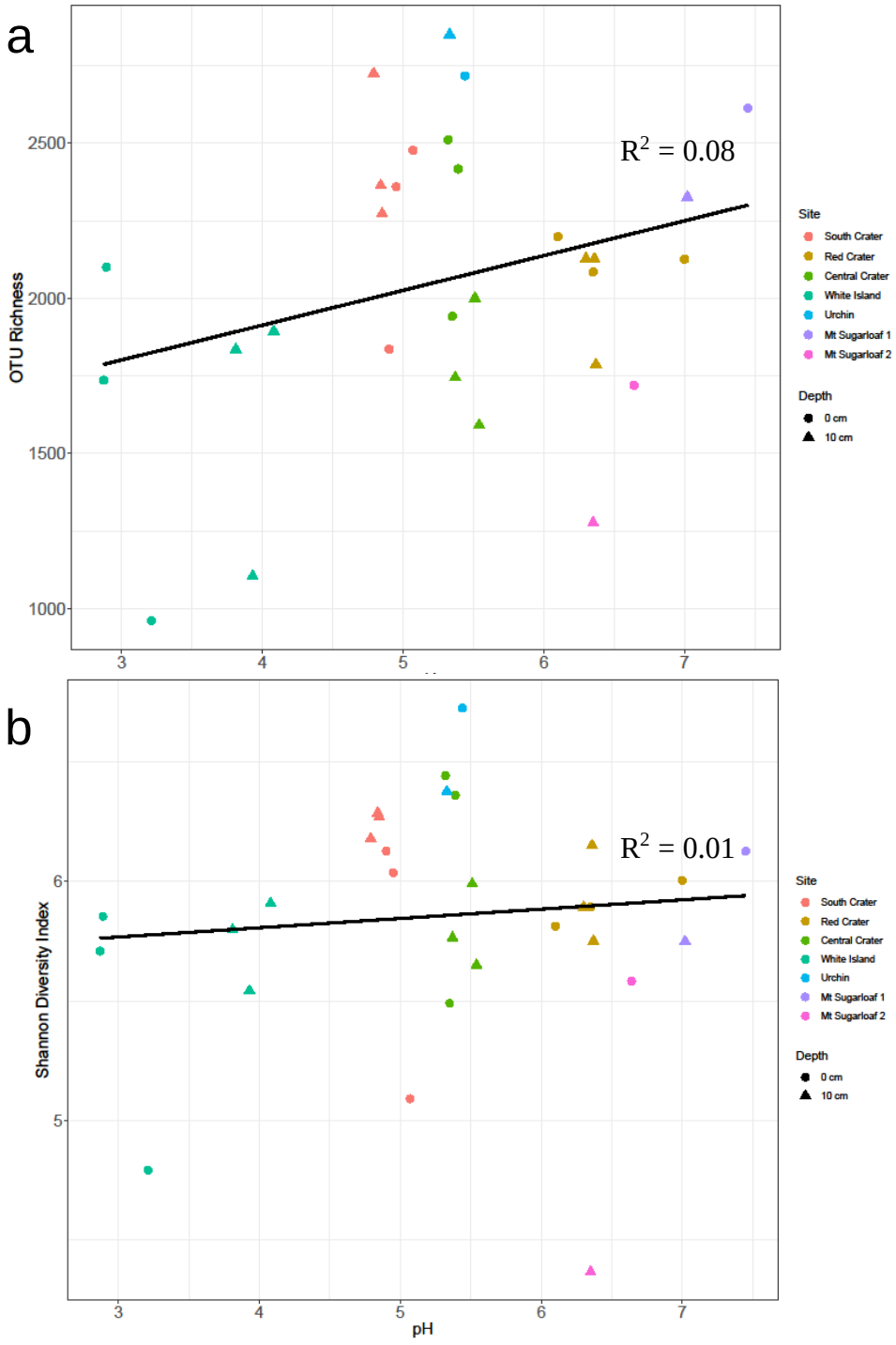

Figure 16: Comparison of pH against alpha diversity indices. a, OTU richness (Observed OTUs) and **b**, evenness (Shannon diversity index, H) of samples are plotted against soil pH. Samples are coloured by site and shaped by depth. Linear regression was applied to these relationships (Observed: $R^2 = 0.08$, $p = 0.12$; Shannon: $R^2 = 0.01$, $p = 0.61$).

Next, to see the compositional differences between the sample sites (beta diversity), Bray-Curtis dissimilarity was calculated. To visualise Bray-Curtis dissimilarity in space, a NMDS ordination is used (Figure 17). This depicts the clustering (or similarity) of sites and depth using the values from Bray-Curtis dissimilarity, Figure 18 shows that samples from the same geographical sites tend to group in space, meaning they have similar beta diversity values (in other words, similar community composition). Sites also tend to cluster based on depth (shown by polygons). Overall, beta diversity is similar between sites indicating the ratio of diversity at each site is similar. All statistical methods are shown in section 2.2.5.

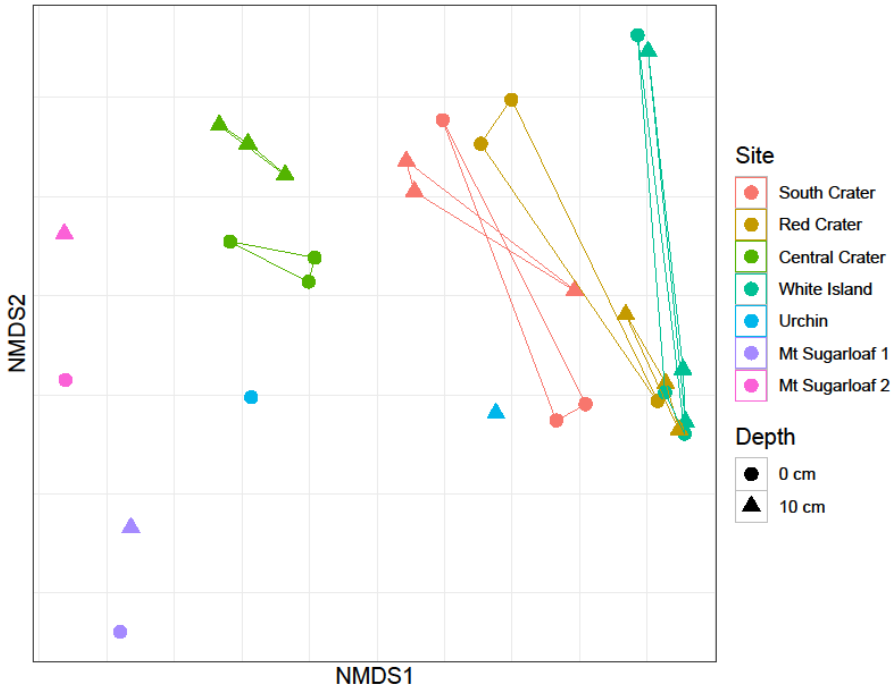

Figure 17: Beta diversity of all soil sample communities. Diversity was calculated using Bray-Curtis dissimilarity and non-metric multidimensional scaling (NMDS, stress = 0.08) was used to ordinate samples in space. No filtering of OTUs was applied prior to calculating Bray-Curtis dissimilarities. Samples are coloured by site and shaped by depth (0 and 10 cm). Samples are also clustered by site using polygons (where n > 3).

To determine the influence of individual metadata or physicochemical variables on the beta diversity of the soil samples, a constrained correspondence analysis (CCA) using a weighted linear regression was undertaken (Figure 18) and all continuous physicochemistry were added to the model in order of highest to lowest Mantel result (Supplementary Table 7); the significant constraining variables lithium, total organic carbon, zinc, manganese, iron, altitude, pH, sodium, magnesium, air temperature, total carbon, cobalt, potassium, silica and soil temperature. Overall, 75.3% of the beta diversity can be attributed to these physicochemical variables. Conversely, if only looking at a CCA of the volcanic sites without controls (e.g. South Crater, Red Crater, Central Crater of Mt. Tongariro and White Island; Figure 19), the most significant constraining variables were lithium, total organic carbon, zinc, manganese, iron, altitude, pH, sodium, magnesium, air temperature and soil temperature. The proportion of variance constrained by these variables was 57.4% (Figure 19).

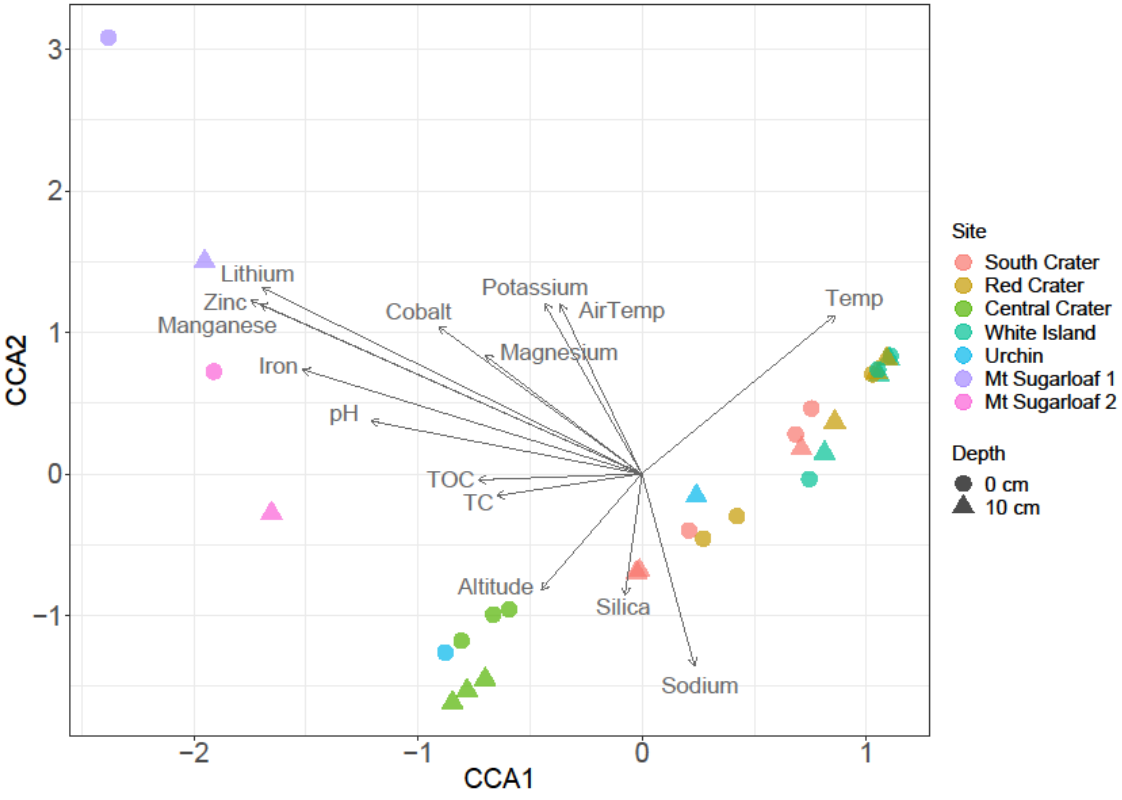

Figure 18: Constrained correspondence analysis (CCA) of beta diversity with significant physicochemistry. Beta diversity was calculated using Bray-Curtis dissimilarities, and all continuous physicochemistry were added to the model in order of highest to lowest Mantel result (except ammonium, caesium and phosphorus due to missing values). Constraining variables, in order of significance, were lithium, total organic carbon (TOC), zinc, manganese, iron, altitude, pH, sodium, magnesium, air temperature, total carbon (TC), cobalt, potassium, silica and soil temperature (Temp). The area and amount of influence are depicted by the

direction and length of the arrows in the ordination. Samples are displayed by site in colour and depth in shape.

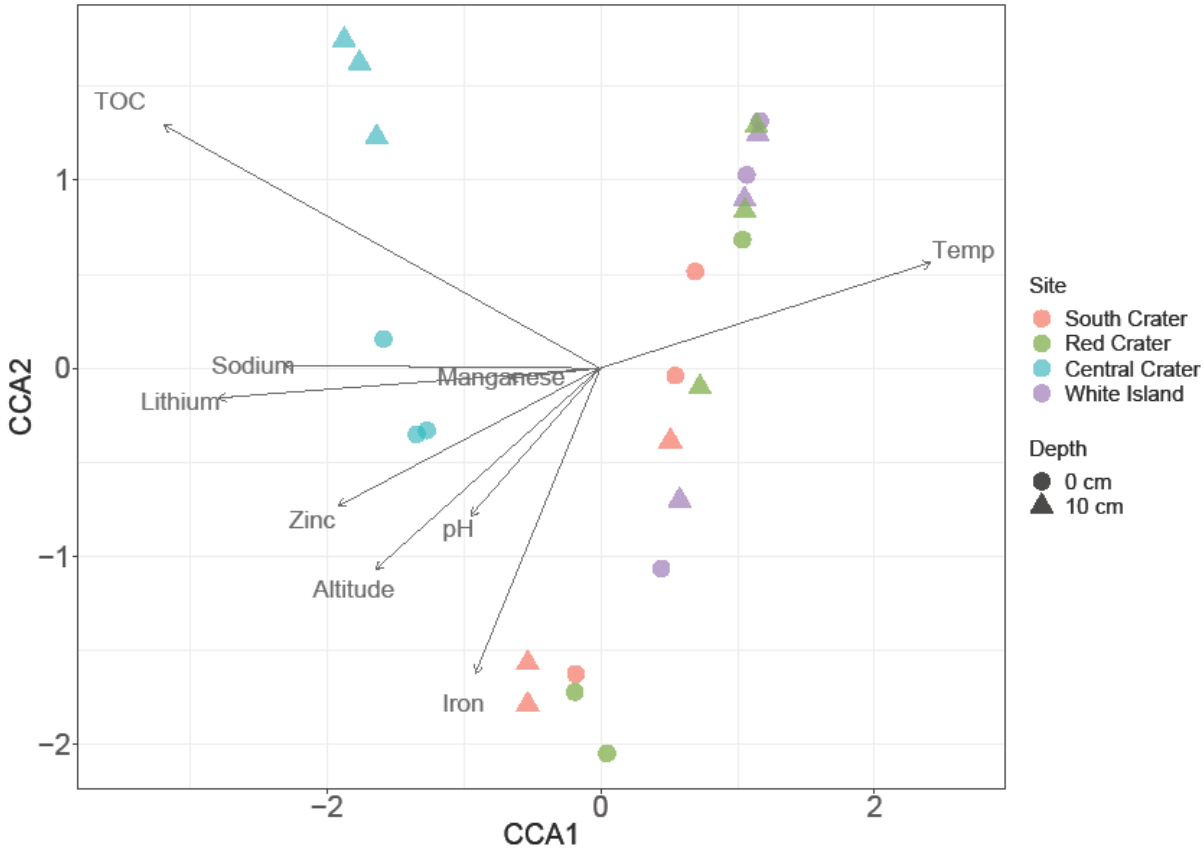

Figure 19: Constrained correspondence analysis (CCA) of beta diversity with significant physicochemistry, without control sites. Beta diversity was calculated using Bray-Curtis dissimilarities, and control sites from Mt. Urchin and Mt. Sugarloaf were removed. All continuous physicochemistry were added to the model in order of highest to lowest Mantel result. Ammonium, boron, caesium and phosphorus had no effect, and were removed from the final model. Constraining variables, in order of significance, were lithium, total organic carbon, zinc, manganese, iron, altitude, pH, sodium, magnesium, air temperature and soil temperature. The proportion of variance constrained by these variables was 57.4%. The area and amount of influence are depicted by the direction and length of the arrows in the ordination. Samples are displayed by site in colour and depth in shape.

2.4 Discussion: Microbial Ecology and Biogeography

A 2010 biogeographical study of high-alpine bacteria showed that spatial distance between microbial communities less accurately predicted community assembly compared to biogeochemical and habitat parameters. It explained that habitat predicts which phyla will be dominant in a system[49]. For Mt. Tongariro and the controls, to see whether physicochemistry was different between the sites, and whether these physicochemical parameters had a significant relationship to both alpha and beta diversity of the communities, statistical analyses were employed in tandem. To see which physicochemical parameters were significantly correlated to alpha diversity, a linear regression comparison of 28 physicochemical parameters was performed. This confirmed silica as the most significant factor influencing Shannon diversity (27.24%, linear regression: *P* 0.0031; Supplementary Table 5). Additionally, multiple linear regression showed silica, calcium and nickel, had significant contributions to alpha diversity (Supplementary Table 6). Cumulatively, physicochemistry accounted for 28% of the alpha diversity observed using this model (Adjusted $R^2 = 0.2875$). These results could be circumstantial, as the soils of Mt. Tongariro are composed predominantly of andesitic rock, which are approximately 63% silica, in the form of silicon dioxide (SiO_2)[174]. Both calculated alpha diversity indices (OTU richness and Shannon's diversity index) indicate physicochemistry has a statistically relevant influence on alpha diversity at the sample site; Shannon diversity index by site (Kruskal-Wallis H = 12.546, p-value = 0.05084) and observed richness by site (Kruskal-Wallis H = 18.039, p-value = 0.006136). However, sample depth had no significant effect on Shannon diversity index (Kruskal-Wallis H = 0.003871, p-value = 0.9504) or observed richness (Kruskal-Wallis H = 0.52688, p-value = 0.4679). In contrast, soil pH or temperature had little influence of alpha diversity indices (Figure 15, 16).

Bray-Curtis dissimilarity was used to show the diversity between sites. Samples from the same geographical sites, and depths, tend to group in space (shown by polygons; Figure 17), indicating they have similar community composition. When constraining beta diversity values against physicochemistry using the CCA model, the most significant constraining variables were lithium, total organic carbon, zinc, manganese, iron, altitude, pH, sodium, magnesium, air temperature and soil

temperature. The proportion of variance in Mt. Tongariro samples constrained by these variables was 57.4% (Figure 19). Using the CCA model, sample sites tend to cluster in space, indicating individual sites have similar diversity, and are affected by the same physicochemical parameters, which corroborates the 2010 study where physicochemistry and habitat are shown as more accurate predictors of community composition than spatial distance[49].

To test whether the diversity and abundance of taxa were comparable to other volcanic, arid or extreme terrestrial environments, analysis of the top phylotypes was done. Arid soil types from cold deserts and high elevation volcanic soils have often been connected with under-dispersed communities dominated by a limited number of phyla[1,2,4]. This is not the case for Tongariro National Park, where the average soil moisture is 16% (w/w) (Supplementary Table 9) and the phylotype diversity is comparable to most soil types worldwide[175,176]. The top phyla observed across all samples were Proteobacteria (20.21%), Chloroflexi (18.17%), Verrucomicrobia (10.54%), Acidobacteria (7.81%), Thaumarchaeota (6.24%), Actinobacteria (5.13%), Planctomycetes (4.82%), Crenarchaeota (4.62%), Deinococcus-Thermus (3.40%), WPS-2 (2.75%), Cyanobacteria (2.58%) and Armatimonadetes (2.34%) (Figure 10; Supplementary Table 3). These phyla are commonly detected in terrestrial environments, especially in elevated, oligotrophic, cold-fumarole volcanic soils[2]. Of the phyla present in the collected samples, the phyla Cyanobacteria, Actinobacteria, Chloroflexi and Deinococcus-Thermus belong to the Terrabacteria superphylum, which harbour adaptations to extreme environments (for example, strains within these phyla can withstand desiccation, ultraviolet radiation, and high salinity) and/or oxygenic photosynthesis, all of which can allow these taxa to persist and colonise a diversity of challenging terrestrial surfaces[42]. With soil moisture in mind, and knowledge that physicochemistry predicts the microbial assemblage[49], it is not surprising then that Mt. Tongariro (~1600-1800 m.a.s.l.; Mt. Urchin ~1300 m.a.s.l.; White Island ~17 m.a.s.l.; Mt. Sugarloaf ~1160 m.a.s.l.) displayed microbial abundance profiles comparable to many terrestrial environments worldwide (Figure 10,[175]). Results from analysis of microbial richness and diversity across all TVZ and control samples showed overall observed richness (961-2,850 phylotypes; mean = 2,060.6) and Shannon diversity index (4.4-6.7; mean = 5.86). This diversity is lower than phylotype richness in various soils ranging from Antarctic cold deserts, temperate and tropical

forests, tundra and prairie grasslands, with richness ranging from 4000-12,000 phylotypes[176], yet on par with the lower end of the spectrum for phylotypes in soils from Central Park, New York (2000-10,000 phylotypes)[175]. Additionally, our TVZ volcanic soils diversity was greater than recent reports of the species richness and diversity of geothermal springs located within the TVZ (49–2,997 phylotypes, mean = 386)[52]. These findings are consistent with observations that OTU richness is generally reduced in aquatic environments compared to terrestrial environments[145].

Physicochemical analysis and microbial abundance results from Mt. Tongariro contrast results from arid, high-elevation alpine Atacama and Antarctic Dry Valley soils ([1,4], and see Section 1.1.4). The major differences are the presence of Cyanobacteria and availability of water. Oligotrophic and arid soil communities from in Atacama Desert and Antarctic Dry Valleys lack Cyanobacteria and had no detectable eukaryotic phototrophs[1,4,38]. Interestingly, McMurdo Dry Valleys have no detectable Cyanobacteria in the soil, but endolith, chasmolith and hypolith communities are dominated by Cyanobacteria (*Chroococcidiopsis* sp. or Oscillatoriales), present where moisture is detectable in rock pores[143]. Differences in cyanobacterial taxa encountered may reflect their adaptation to moisture availability in oligotrophic systems. Similarly, soils of Robinson Ridge in the Antarctic Dry Valleys contain only 0.28% mean relative abundance of Cyanobacteria[1]. Mt. Tongariro soils, in contrast had 10-fold the amount of Cyanobacteria (2.8%). This difference in Cyanobacteria is likely due to the increased soil moisture; Mt. Tongariro soils contain an average of 16% (v/v) moisture content (Supplementary Table 9). In support, the abundance of cyanobacteria in rock endoliths of the Atacama Desert is credited to atmospheric moisture, where microbial populations in rock pores rely on oxygenic photosynthesising Cyanobacteria for primary production[177], and eukaryotic phototrophy is limited due to soil moisture and atmospheric moisture being below the calculated 1.3% limit for eukaryotic life[38]. In contrast, there has been no dry limit detected for bacterial cells[38,177]. Though there is no dry limit detected for bacterial cells, there is a correlation between moisture availability, either in the soil or atmosphere and the presence of Cyanobacteria in volcanic soils and rocks[143,177].

The presence of moisture in Mt. Tongariro could be contributing to these observed differences in the microbial populations between Mt. Tongariro and those of dry volcanic Atacama Desert soils and Antarctic soils. Conceivably, the mode of primary production in hyperarid soils would be trace-gas oxidising taxa from candidate phyla WPS-2, AD3, Chloroflexi and Actinobacteria, which are commonly detected in oligotrophic, bare soils and hyperarid soils, especially where the system is devoid of any detectable phototrophs. However, Cyanobacteria when present are likely functioning as primary producers in the system. It has been noted that Cyanobacteria secrete hygroscopic extracellular polymeric substance (EPS) to combat desiccation which is their mode of adaptation to inhabit rock surfaces in hyperarid deserts such as Atacama Desert and Antarctic dry valley rocks[143]. The mode of primary production by oxygenic phototrophic Cyanobacteria is by carbon input into the system which is usually accompanied by nitrogen fixation by other taxa, and possibly supported by atmospheric trace-gas oxidising taxa. Interestingly, at the genus level, Mt. Tongariro volcanic and bare control soils contain taxa within the genera *Pseudonocardia*, from the phylum Actinobacteria, with a relative abundance of 2.81% with a maximum of 9.36%. Actinobacteria lineages from the genus *Pseudonocardia* are known trace gas (CO and H_2) oxidisers. *Pseudonocardia* encode group 1h [NiFe] hydrogenase enzymes which are capable of oxidising atmospheric levels of H_2[37,178,179] and also encode type I [MoCu] carbon monoxide dehydrogenase, a family of enzymes responsible for the oxidation of atmospheric CO[1].

Due to the high proportion of trace-gas oxidising taxa in Mt. Tongariro and the control soils from Mt. Urchin and Mt. Sugarloaf, a microcosm study of *in situ* trace gas oxidation of these soils was completed (see Chapter 3 below).

2.5 Conclusions: Microbial Ecology and Biogeography

The composition and diversity of soils from Mt. Tongariro and the control soils were comparable to many arid, oligotrophic and temperate terrestrial systems on Earth. Soil moisture seems to influence the presence of Cyanobacteria in the soil which is corroborated by previous literature. Additionally, Mt. Tongariro soils harbour many known trace-gas oxidising phyla including Actinobacteria, Chloroflexi, WPS-2 and AD3. Samples tend to cluster by site, or habitat, as expected. Interestingly, pH and temperature did not significantly influence alpha diversity, showing the driver of diversity must be other physicochemical parameters. The physicochemistry accounted for up to 28% of the alpha diversity using the multi linear regression model and up to 57% of the beta diversity using a CCA model. In summation, the microbial communities were comparable to worldwide soils, yet had the presence of major taxa inhabiting hyperarid volcanic and cold desert soil types and these taxa are known oxidisers of trace gas.

Chapter 3: Microbial Trace Gas Oxidation

3.1 Introduction

In situ, thermophilic isolates and soil microcosms have been shown to persist throughout periods of carbon limitation (starvation) by oxidising trace levels of hydrogen and carbon monoxide[1,44,45]. However, the mechanisms thermophilic taxa employ to persist within environments that have lost aspects of extremophilic classification, such as dormant volcanic and geothermal systems remain unresolved. Environmental sampling of non-extreme (i.e. moderate temperature) environments continues to reveal the presence of thermophilic taxa despite an absence of heating necessary to accommodate thermophilic growth[8]. These findings beg the question; are the mechanisms of persistence due to sporulation, a slowing of metabolic activity towards dormancy or are the traces (DNA sequence OTUs) of thermophilic taxa detected in cool temperature (i.e. non-thermophilic) merely a relic (e.g. adsorbed to soil as a fossil to their previous environment)?

The primary aim of Chapter 3 is to determine if the oxidation of trace gases is occurring in soils from Mt. Tongariro. Do microbial taxa persist in these oligotrophic volcanic soils by using H_2 and CO as energy *in lieu* of traditional primary production? Microcosms amended with trace gases into their headspaces were incubated at 4 °C and 60 °C to investigate whether thermophilic taxa detected (Supplementary Table 4) were persisting in these environments. As many putative methane oxidisers were present in these soil sites (according to our DNA sequence data – Supplementary Table 4), CH_4 oxidation was also assessed for each sample site at 4 °C and 60 °C. Microcosm headspace gas concentrations from volcanic soils of Mt. Tongariro – South, Red and Central Craters were tracked over time. Also, volcanic soils from White Island (Whakaari), alpine and bare soils from Mt. Urchin (a non-volcanic mountain range adjacent to the Taupō Volcanic Zone (TVZ)) and bare elevated soils from Mt. Sugarloaf from New Zealand's South Island were used as controls for this experiment. Trace gas oxidation (CO, H_2, CH_4) in these soil microcosms was assessed using gas chromatography.

A secondary aim of Chapter 3 was to test whether the oxidation of H_2 and CO occurs at thermophilic temperatures (60 °C) by using the same soil microcosms and control sites mentioned above. The presence of genes encoding high-affinity enzymes

responsible for the oxidation of H_2 and CO at atmospherically relevant concentrations in sample soils was screened for via PCR and primers targeting these specific gene sequences.

3.2 Methods

3.2.1 Gas Chromatography

Gas chromatography was used to determine whether the oxidation of atmospheric gases (H_2, CO, and CH_4) occurs in alpine, volcanic and non-volcanic bare soil microbial communities at both 4 °C and thermophilic settings (60 °C). For each sample site, biological soil triplicates were taken aseptically from the surface, 0 cm and 10 cm depth from South, Red and Central Craters of Mt. Tongariro, Tongariro National Park; biological triplicates were retrieved from White Island and technical triplicates were taken from Mt. Urchin and Mt. Sugarloaf, both at 0 cm and 10 cm depth. To construct soil microcosms, soil samples of 1 g were aseptically added into sterile 114 ml serum bottles (Supelco, Sigma-Aldrich, USA) and sealed with butyl rubber stoppers (Chemglass Life Science, USA).

To the soil microcosms, H_2 and CO gases were added to achieve an initial headspace concentration of 100 parts per million by volume (p.p.m.v.) and 20 p.p.m.v., respectively, as described previously[44,145]. To compliment observations of H_2 and CO oxidation, microcosms amended with 100 p.p.m.v CH_4 were also prepared. These headspace concentrations were selected to reflect current tropospheric mixing ratios of 0.53 p.p.m.v. for H_2, 0.1 p.p.m.v. for CO [145] and 1.8 p.p.m.v. for CH_4[180]. It should be noted that large variations in partial concentrations of trace gas are common in urban and geothermal areas.

Microcosm headspaces were sampled (1 ml) using a gas-tight syringe (SGE Gas Tight Syringe, Sigma-Aldrich, USA). H_2 and CO were measured using a PP1 Peak Performer 1 Gas Analyser (Peak Laboratories, Mountain View, CA, USA) equipped with a reducing compound photometer (RCP: H_2, CO, 100 µl sampling loop), or flame ionizing detector (FID: CH_4, 25 µl sampling loop), Unibeads 1S 60/80 column, and Molecular Sieve 13X 60/80 column. Heat-killed soils (1 g; 121 °C, 15 p.s.i., 20 min) were used as negative controls and were analysed in parallel to regular headspace sampling to ensure that the observed losses of H_2, CO

and CH_4 were the consequences of biological activity. Blank air controls and gas standards were analysed before each sampling run to ensure GC values were precise and accurate. Additionally, for field gas sampling, concentrations of H_2, O_2, N_2, CH_4 and CO_2 were measured using a 490 Micro Gas Chromatograph equipped with a thermal conductivity detector (TCD; Agilent Technologies, USA). Field gas measurements can be found in Supplementary Table 10.

All experiments were repeated at least once. Microcosm gas oxidation experiments were initially incubated at 4 °C. Before the start of a repeated experiment, to ensure there was sufficient oxygen in the headspace, serum bottles were vented, and sterile air was added to the headspace using a 0.22 µm syringe filter (Minisart®, Sartorius, Göttingen, Germany). Following this, atmospheric concentrations of H_2, CO, or CH_4 gases were added aseptically using a 0.22 µm syringe filter as outlined above. Following the completion of 4 °C experiments, microcosms were incubated at 60 °C to investigate the occurrence of trace gas oxidation at thermophilic temperatures. Headspaces were flushed again with sterile air using a 0.22 µm filter and filled to ~100 p.p.m.v. H_2, ~20 p.p.m.v. CO (and ~150 p.p.m.v. CH_4). One of the biological triplicates, and one technical triplicate (which previously oxidised gas to under atmospheric levels) for each site was retained at 4 °C as an experimental control, to ensure a potential lack of observed gas oxidation at 60 °C could not be attributed to oxygen or soil nutrient limitation caused by the first set of experiments.

3.2.2 Data Analysis

Trace-gas oxidation was monitored using gas chromatography methods outlined in Section 3.2.1. Values were entered to an excel spreadsheet with time in hours (h) and value calculated as p.p.m.v. Final graphs were produced using GraphPad Prism v7.03. Below are the final XY scatterplots tracking results of trace gas oxidation in soil microcosms at 4 °C and 60 °C of CO, H_2 and CH_4. Mt. Urchin Track represent North Island control soil (non-volcanic, yet TVZ adjacent) and Mt. Sugarloaf are South Island, non-volcanic controls. Each experiment took between 1500 – 3000 h for complete oxidation of H_2 and CO. All XY scatterplots were produced using technical or biological triplicate values and the data points represent the mean value for each concentration at each time point.

3.2.3 Amplification of genes encoding for group 1h [NiFe]-hydrogenases (*hhyL*) and type 1 [MoCu]-carbon monoxide dehydrogenases (*coxL*)

Targeted PCR was used to confirm the presence of group 1h [NiFe]-hydrogenase large subunit gene (*hhyL*) using primers (NiFe-244f: 5′-GGGATCTGCGGGGACAACCA-3′); and (NiFe568r: 5′-TCTCCCGGGTGTAGCGGCTC-3′)[181]. PCR was also used to confirm the presence of type 1 [MoCu]-carbon monoxide dehydrogenases large subunit gene (*coxL*) primers based on sequences from Robinson Ridge, Antarctica using primers (coxL1288f: 5′-TSKKYACSGGCWSSTA-3′); and (coxL1540r: 5′-TAYGAYWSSGGYRAYTA-3′)[1].

PCR master mixture per reaction contained: 35 µl of PCR H_2O, 5 µl dNTP, 5 µl 10X Taq buffer, 0.5 µl Primer F (0.5µM), 0.5 µl Primer R (0.5µM), 0.5 µl of Taq 5U μl^{-1} and 2.5 µl of template DNA (~50 ng DNA). PCR: Initial denaturation was achieved at 94 °C for 5 min, then 29 PCR cycles consisted of denaturing for 45 s at 94 °C, annealing for 45 s at a temperature gradient of 54 - 58 °C and extension for 1 m 30 s at 72 °C. A final extension was completed for 5 min at 72 °C, using a T100™ Thermal Cycler (Bio-Rad, Hercules, CA, USA). To confirm the identity of PCR products, bands were visualised by gel electrophoresis, using 2% SeaKem® LE Agarose (Lonza, Rockland, ME, USA). Bands of the correct product size (~780 bp *coxL*, ~350 bp *hhyL*) were aseptically cut from the gel and purified using NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel, Düren, Germany). Sanger sequencing was performed on these forward and reverse primer sets (Macrogen, Geumcheon-gu, Seoul)[182].

3.3 Results: PCR, Enzymes and Trace Gas Oxidation at 4 °C

3.3.1 Hydrogenase and Carbon Monoxide Dehydrogenase Detection

Figure 20 shows amplification of Actinobacteria-type 1h [NiFe] hydrogenases using targeted primer sets described in methods 3.2.3. Figure 21 shows amplification of Type I Aerobic CODH. The Sanger sequencing quality and PCR product was poor for Mt. Urchin, White Island and Central Crater and

consequently did not result in reliable final sequences. The sequences with high quality reads were run against the Basic Local Alignment Search Tool (BLAST) to find similarities to other nucleotide sequences in the database[183]. PCR no template controls (negative controls) were run, as expected, the lanes were empty (data not shown). Results show that all the CODH sequences with high quality reads were most similar to the large subunit (coxL) of the Type I [MoCu] carbon monoxide dehydrogenase in uncultured bacterial clones. Four of the five high quality sequence reads for Type 1h [NiFe]- hydrogenase were most similar in sequence to *Streptomyces*; nucleotide sequence similarity range 80-86%. The remaining sequence was closest to an uncultured bacterial clone; sequence similarity 74-100% (Table 2).

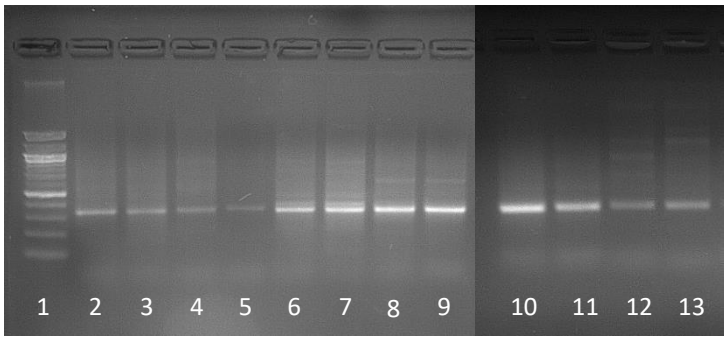

Figure 20: Gel electrophoresis image showing [NiFe]-hydrogenase band amplifications. Row 1, 100 bp DNA Marker Ladder (Zymo Research, USA). Row 2, South Crater 0 cm, Row 3: South Crater 10 cm, Row 4: Red Crater 0 cm, Row 5: Red Crater 10 cm, Row 6: Central Crater 0 cm Row 7: Central Crater 10 cm, Row 8: Mt. Sugarloaf 0 cm, Row 9: Mt. Sugarloaf 10 cm, Row 10: White Island 0 cm, Row 11: White Island 10 cm, Row 12-13: Positive Control (*Pyrinomonas methylaliphatogenes* K22^{T}). Negative controls not shown.

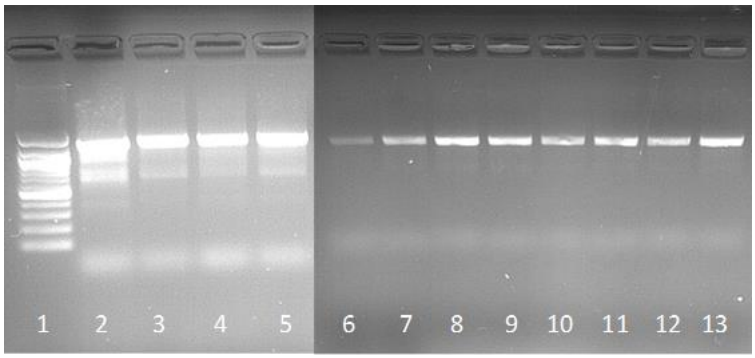

Figure 21: Gel electrophoresis image showing [MoCu]-CODH band amplifications. Row 1, 100 bp DNA Marker Ladder (Zymo Research, USA). Row 2, South Crater 0 cm, Row 3: South Crater 10 cm, Row 4: Red Crater 0 cm, Row 5: Red Crater 10 cm, Row 6: Central Crater 0 cm Row 7: Central Crater 10 cm, Row 8: Mt. Sugarloaf 0 cm, Row 9: Mt. Sugarloaf 10 cm, Row 10: White Island 0 cm, Row 11: White Island 10 cm, Row 12-13: Positive Control (*Thermogemmatispora* sp. T81). Negative controls not shown.

site	*hhyl*	Accession	Identity %
South Crater 0 cm	*Streptomyces atratus* strain SCSIO_ZH16 chromosome, complete genome	CP027306.1	85%
Central Crater 0cm	*Streptomyces* sp. 3124.6 genome assembly, chromosome: I	LT670819.1	86%
Central Crater 10 cm	Uncultured bacterium clone FS6 [NiFe]-hydrogenase large subunit	HM029266.1	84%
Mt Sugarloaf 0 cm	*Streptomyces* sp. TN58, complete genome	CP003777.1	80%
Mt Sugarloaf 10 cm	*Streptomyces scabiei* hhyL gene for [NiFe]-hydrogenase large subunit	AB894417.1	86%
	coxL		
South Crater 10 cm	Uncultured bacterium clone TF-2_115 CODH form I large subunit gene	KJ468337.1	87%
Central Crater 0cm	Uncultured bacterium clone GL13-8_(OMP) CODH form I coxL	KC843700.1	74%
Central Crater 10 cm	Uncultured bacterium clone LD4-9_COX_6 large subunit CODH coxL	KF524338.1	80%
Mt Sugarloaf 0 cm	Uncultured bacterium clone CTE1014 CODH (coxL) gene, partial cds	KT460823.1	100%
Mt Sugarloaf 10 cm	Uncultured bacterium clone FG-1_181 CODH form I large subunit gene	KJ468269.1	77%
site	***hhyl***	**Accession**	**Identity %**
South Crater 0 cm	*Streptomyces atratus* strain SCSIO_ZH16 chromosome, complete genome	CP027306.1	85%
Central Crater 0cm	*Streptomyces* sp. 3124.6 genome assembly, chromosome: I	LT670819.1	86%
Central Crater 10 cm	Uncultured bacterium clone FS6 [NiFe]-hydrogenase large subunit	HM029266.1	84%
Mt Sugarloaf 0 cm	*Streptomyces* sp. TN58, complete genome	CP003777.1	80%
Mt Sugarloaf 10 cm	*Streptomyces scabiei* hhyL gene for [NiFe]-hydrogenase large subunit	AB894417.1	86%
	coxL		
South Crater 10 cm	Uncultured bacterium clone TF-2_115 CODH form I large subunit gene	KJ468337.1	87%
Central Crater 0cm	Uncultured bacterium clone GL13-8_(OMP) CODH form I coxL	KC843700.1	74%
Central Crater 10 cm	Uncultured bacterium clone LD4-9_COX_6 large subunit CODH coxL	KF524338.1	80%
Mt Sugarloaf 0 cm	Uncultured bacterium clone CTE1014 CODH (coxL) gene, partial cds	KT460823.1	100%
Mt Sugarloaf 10 cm	Uncultured bacterium clone FG-1_181 CODH form I large subunit gene	KJ468269.1	77%

Table 2: Nucleotide sequence similarity results based on Type 1h [NiFe]-hydrogenase (*hhyl*) and Type I [MoCu] carbon monoxide dehydrogenase (*coxL*) sequences.

3.3.2 Trace-gas oxidation at 4 °C

Bare soils from Mt. Urchin and Mt. Tongariro (South Crater, Red Crater and Central Crater) at depth of 0 cm and 10 cm were all observed to oxidise H_2 to under atmospheric concentrations (<0.5 p.p.m.v.) (Figure 22, a). Complete oxidation of H_2 was more rapid in Central Crater (0 cm), Mt. Urchin (0 cm) and Red Crater (0 cm) soils (Figure 22, a) and Mt. Sugarloaf (0 cm) (Figure 25, a) compared to the corresponding sites at depth (10 cm), which all took less than 500 hours for oxidation to under tropospheric H_2 concentrations.

Notably, overall for the surface depth (0 cm) at sites at Mt. Urchin, Red Crater, Central Crater and Mt. Sugarloaf, complete oxidation was more rapid at the surface (0 cm), compared to depth (10 cm). South Crater was the exception, in which the soils at depth oxidised H_2 more rapidly. Similarly, bare soils from Mt. Urchin and Mt. Tongariro (South Crater, Red Crater and Central Crater) at depths of 0 cm and 10 cm were all able to oxidise CO to under concentrations (<0.1 p.p.m.v.), except South Crater (0 cm) which could oxidise CO down to 1 p.p.m.v. but not to under atmospheric concentrations (<0.1 p.p.m.v.) (Figure 22, b). Red Crater (10 cm) and Mt. Urchin (0 cm and 10 cm) samples displayed the most rapid complete oxidation of CO to below atmospheric concentration (< 500 hours). Heat kill controls for both H_2 and CO oxidation maintained stable partial pressures throughout experimentation, indicating the oxidation of H_2 and CO must be from biological processes. Similarly, soils at depth (10 cm) from Mt. Sugarloaf (Figure 24, b) oxidised CO and both 0 cm and 10 cm depth were capable of oxidising trace gas to under tropospheric partial pressures (<0.1 p.p.m.v.). Finally, soil microcosms from White Island did not oxidise gas (H_2, CO or CH_4) (Figure 24, a). Nor was CH_4 oxidation observed in any of the soils tested (Mt. Urchin, South Crater, Red Crater, Central Crater (Figure 23, b) or Mt. Sugarloaf (Figure 24, c)).

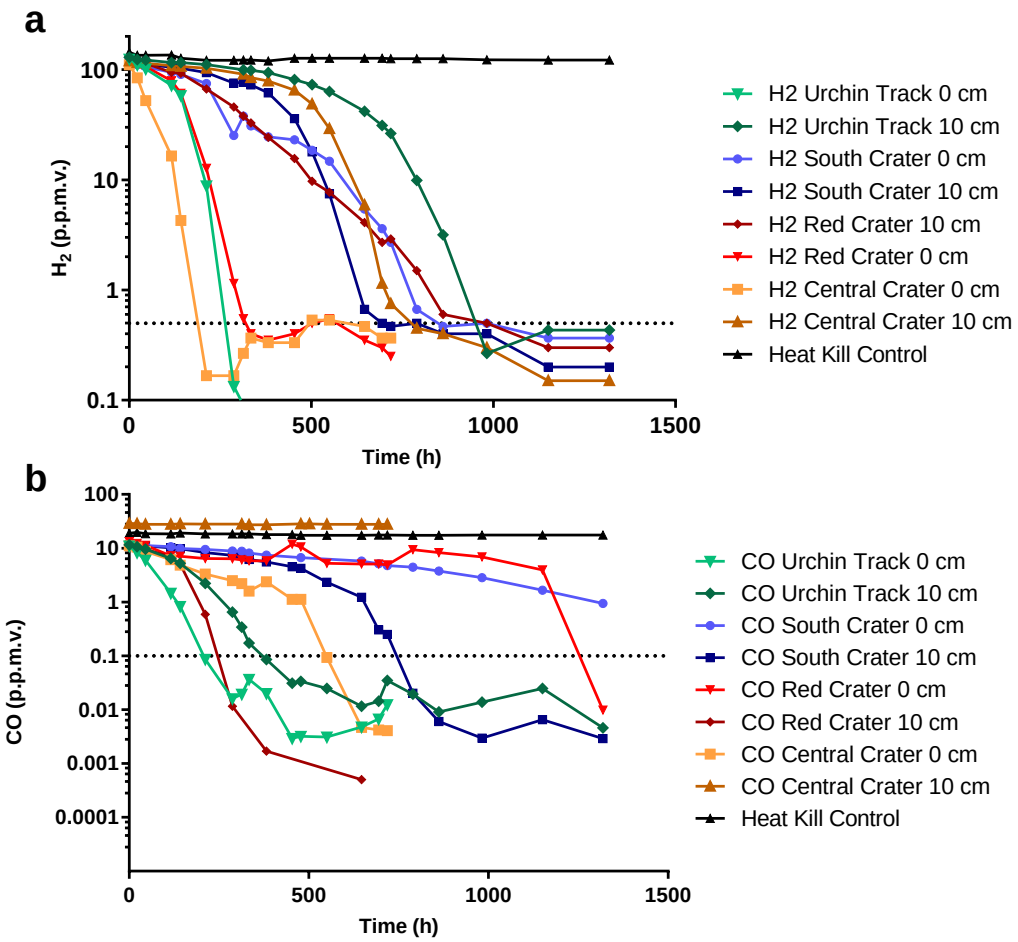

Figure 22: Trace gas oxidation of H_2 and CO of Mt. Tongariro (South, Red and Central Craters) and Mt. Urchin soil microcosms. Gas chromatography measurements of **a,** oxidation of atmospheric H_2 (mixing ratio, 0.5 parts per million by volume (p.p.m.v.) at 4 °C **b,** oxidation of atmospheric CO (mixing ratio, 0.1 parts per million by volume (p.p.m.v.) at 4 °C.

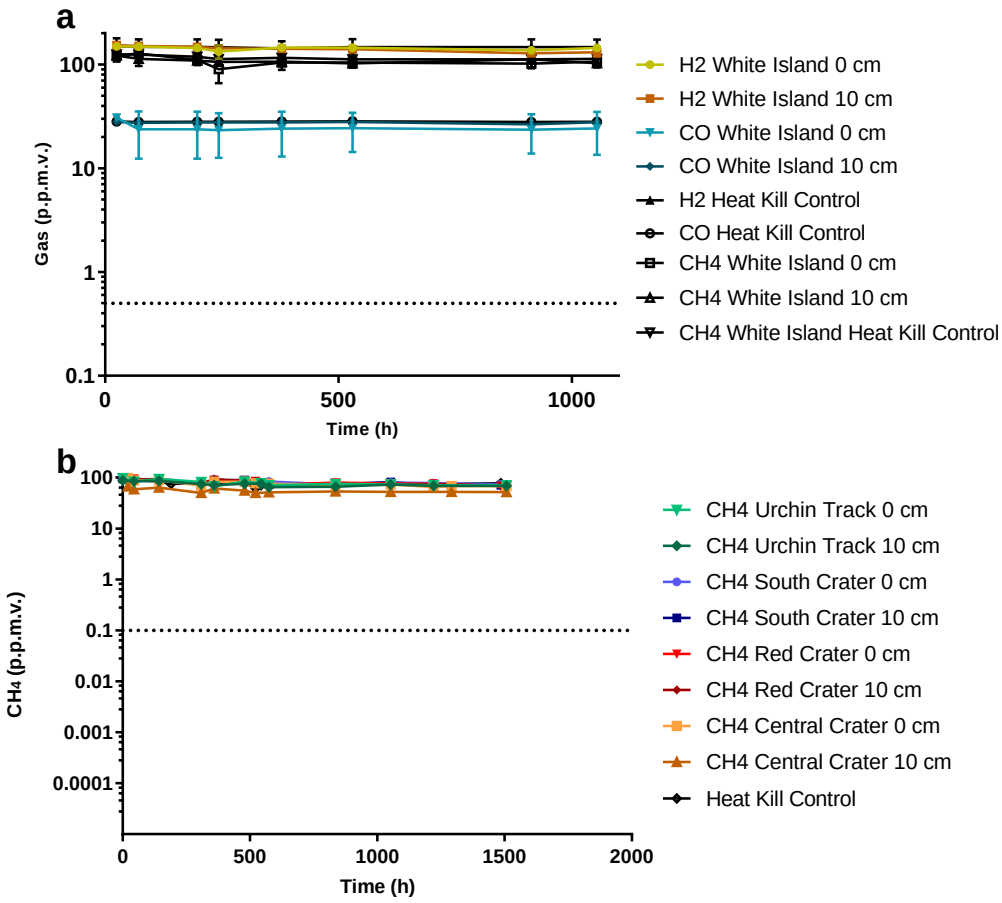

Figure 23: Trace gas oxidation of H_2, CO and CH_4 in volcanic (White Island, Mt Tongariro) and volcanic-adjacent (Mt Urchin) soil microcosms at 4 °C. Gas chromatography measurements of **a,** oxidation of atmospheric H_2 and CO (H_2 mixing ratio, 0.5 p.p.m.v., CO mixing ratio, 0.1 p.p.m.v.) at 4 °C. **b,** oxidation of methane gas (CH_4) (CH_4 mixing ratio 1.8 p.p.m.v.) at 4°C).

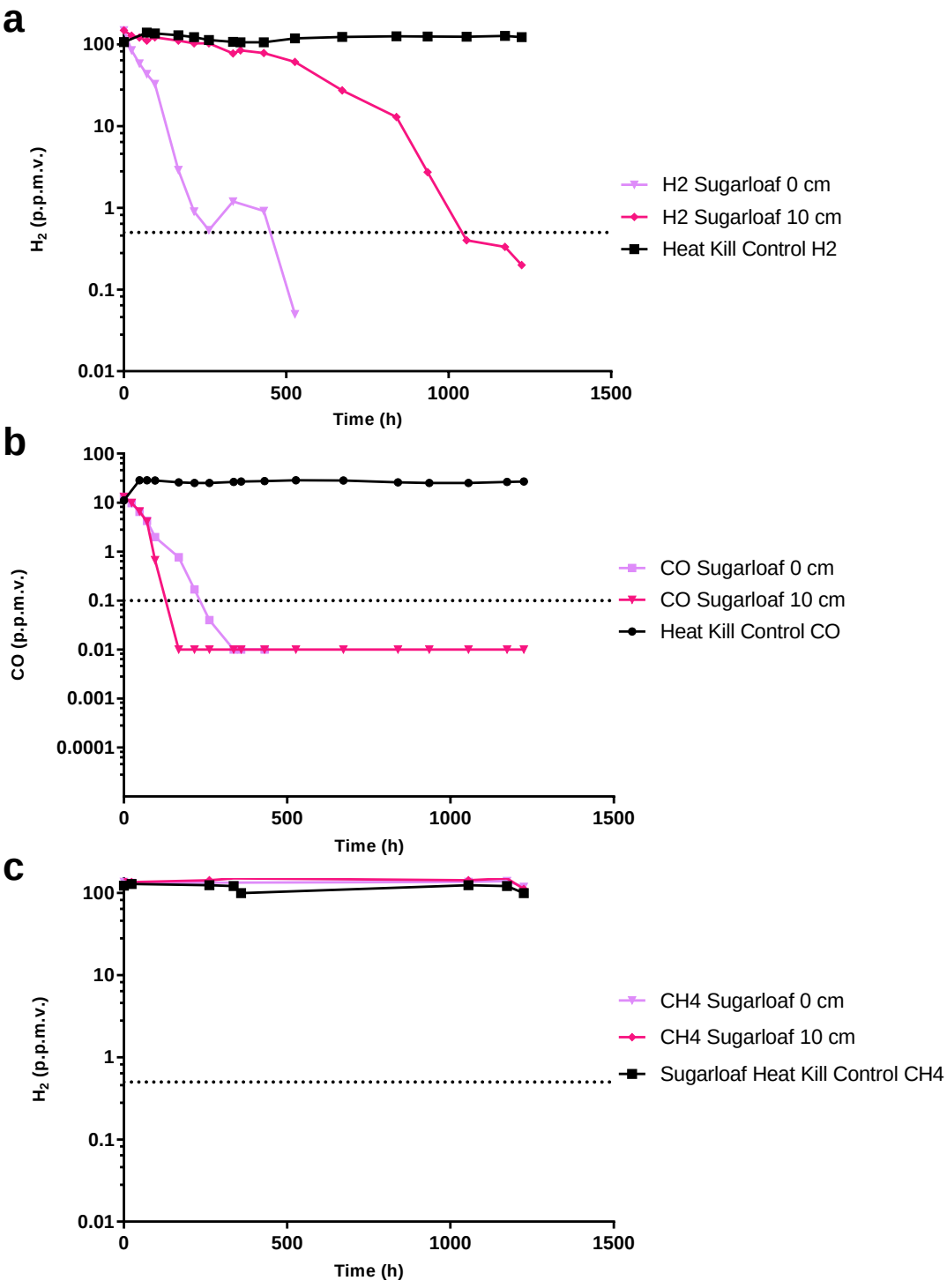

Figure 24: Trace gas oxidation of H_2, CO and CH_4 in Mt Sugarloaf soil microcosms. Gas chromatography measurements of oxidation of **a,** atmospheric H_2 (mixing ratio, 0.5 parts per million by volume (p.p.m.v.) **b,** atmospheric CO (mixing ratio, 0.1 p.p.m.v.) and c, atmospheric CH_4, (mixing ratio, 1.8 p.p.m.v.) at 4 °C.

3.3.3 Trace-Gas-Oxidation at 60 °C

Trace gas oxidation was investigated at 60 °C for the same soil microcosms used for the 4 °C trace gas oxidation experiments. One of the technical or biological triplicates was retained at 4 °C to ensure oxidation (or lack thereof) was the consequence of biological processes, not due to a lack of oxygen for aerobic oxidation of H_2 or CO. Figures 25-29 show the results of the thermophilic trace gas oxidation experiments. As expected, trace gas oxidation of H_2 and CO at 4 °C was observed, and sterile (heat killed) soils retained their original gas partial pressures (i.e. they are not oxidising trace gas).

Complete oxidation of H_2 in thermophilic (60 °C) soil microcosms was observed for Mt. Urchin (0 and 10 cm) and South Crater (10 cm) samples. These sites were capable of complete H_2 oxidation to under tropospheric partial pressures (0.5 p.p.m.v.) in under 500 hours (Figure 25). South Crater surface soils were capable of oxidising H_2 from 100 p.p.m.v. to ~ 10 p.p.m.v. but were not able to oxidise H_2 to less than tropospheric levels. Thermophilic (60 °C) soil microcosms at Red Crater (0 cm and 10 cm) and Central Crater (0 cm and 10 cm) were capable of oxidation of H_2 from 100 p.p.m.v. to between 1 - 10 p.p.m.v. yet were not capable complete oxidation of CO at 60 °C was only observed in Mt. Urchin (Figure 27, a) soils. Soils at Red Crater (0 cm) and Central Crater (0 cm) (Figure 29) were both capable of oxidising CO to ~ 1 p.p.m.v., but not to under atmospheric concentrations (< 0.1 p.p.m.v.). Oxidation of CO at Red Crater (10 cm) and Central of further oxidation to less than atmospheric concentrations (Figure 26). Similarly, Crater (10 cm) however is undetectable above background variation when compared to the natural variation in heat kill control values (Figure 28). Consistent with results from all sites at 4 °C (Mt. Urchin, Mt. Tongariro, White Island), no CH_4 oxidation in thermophilic (60 °C) soil microcosms (Figure 29) and no CO or H_2 oxidation for White Island samples was observed (Figure 29, a).

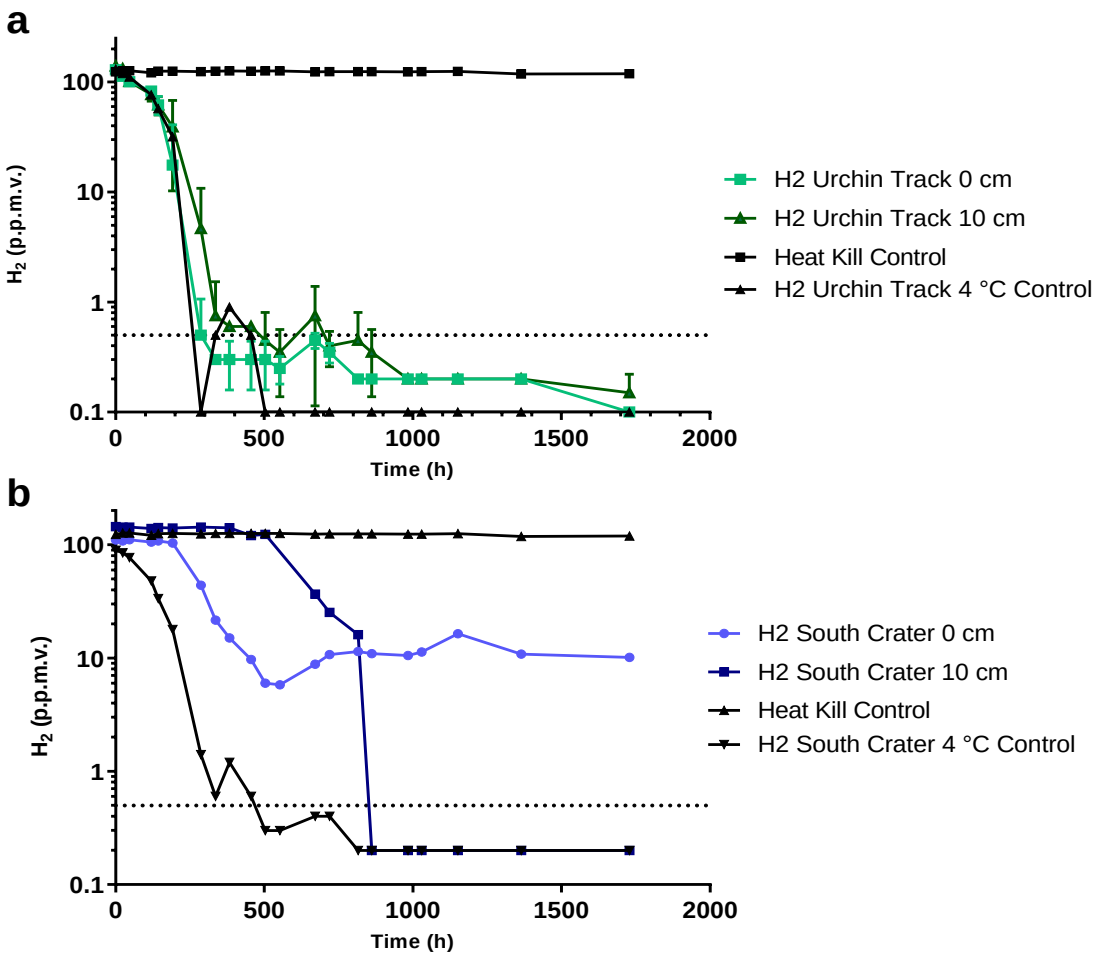

Figure 25: Thermophilic trace gas oxidation of H_2 in volcanic (Mt. Tongariro) and volcanic-adjacent (Mt. Urchin) soil microcosms. Gas chromatography measurements of 60 °C incubations tracking oxidation of atmospheric H_2 (mixing ratio, 0.5 parts per million by volume (p.p.m.v.) at **a,** Mt. Urchin Track, Kaimanawa Forest Park and **b,** South Crater. Mt. Tongariro.

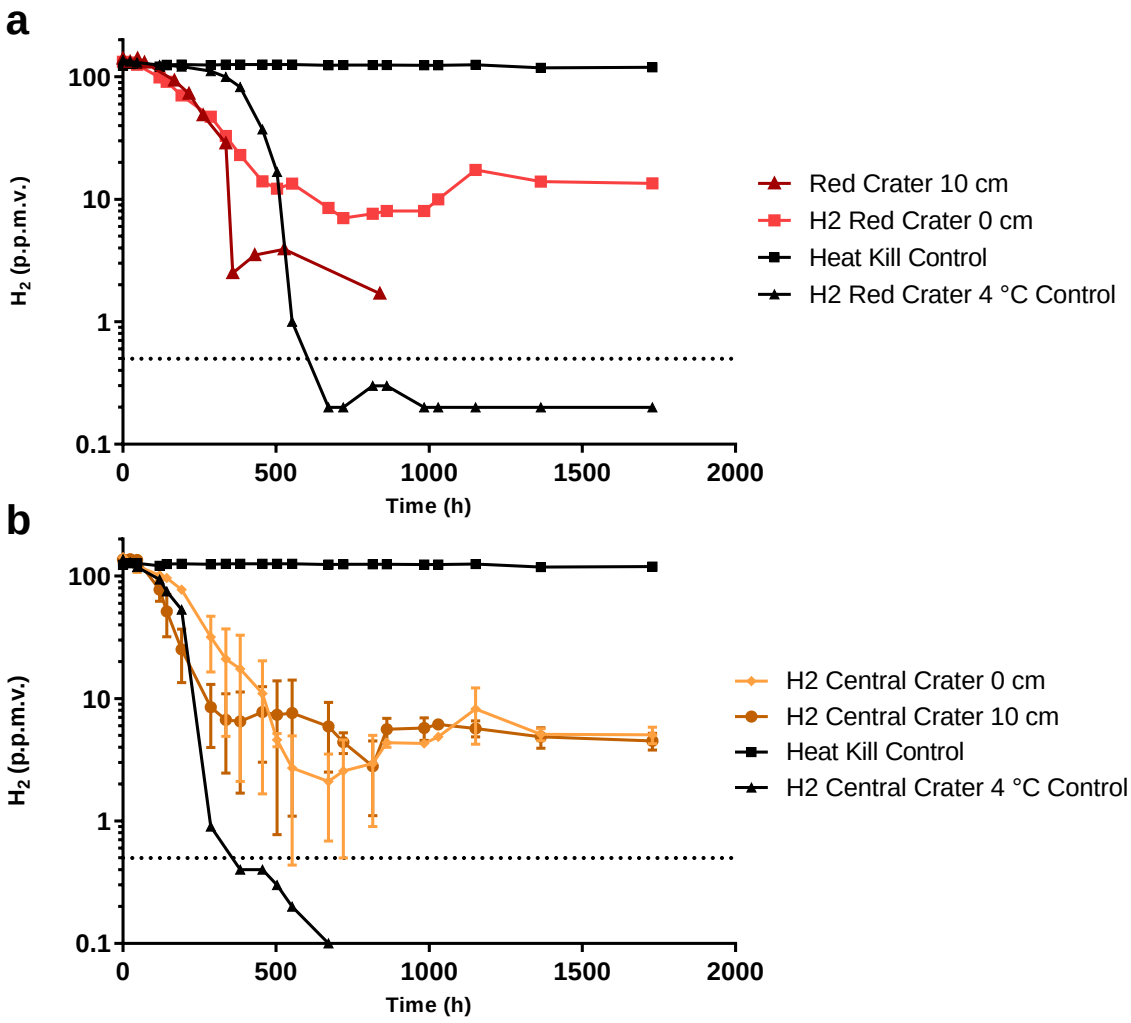

Figure 26: Thermophilic trace gas oxidation of H_2 in volcanic soil microcosms. Gas chromatography measurements of 60 °C incubations tracking oxidation of atmospheric H_2 (mixing ratio, 0.5 parts per million by volume (p.p.m.v.) at **a,** Red Crater, Mt. Tongariro and **b,** Central Crater, Mt. Tongariro.

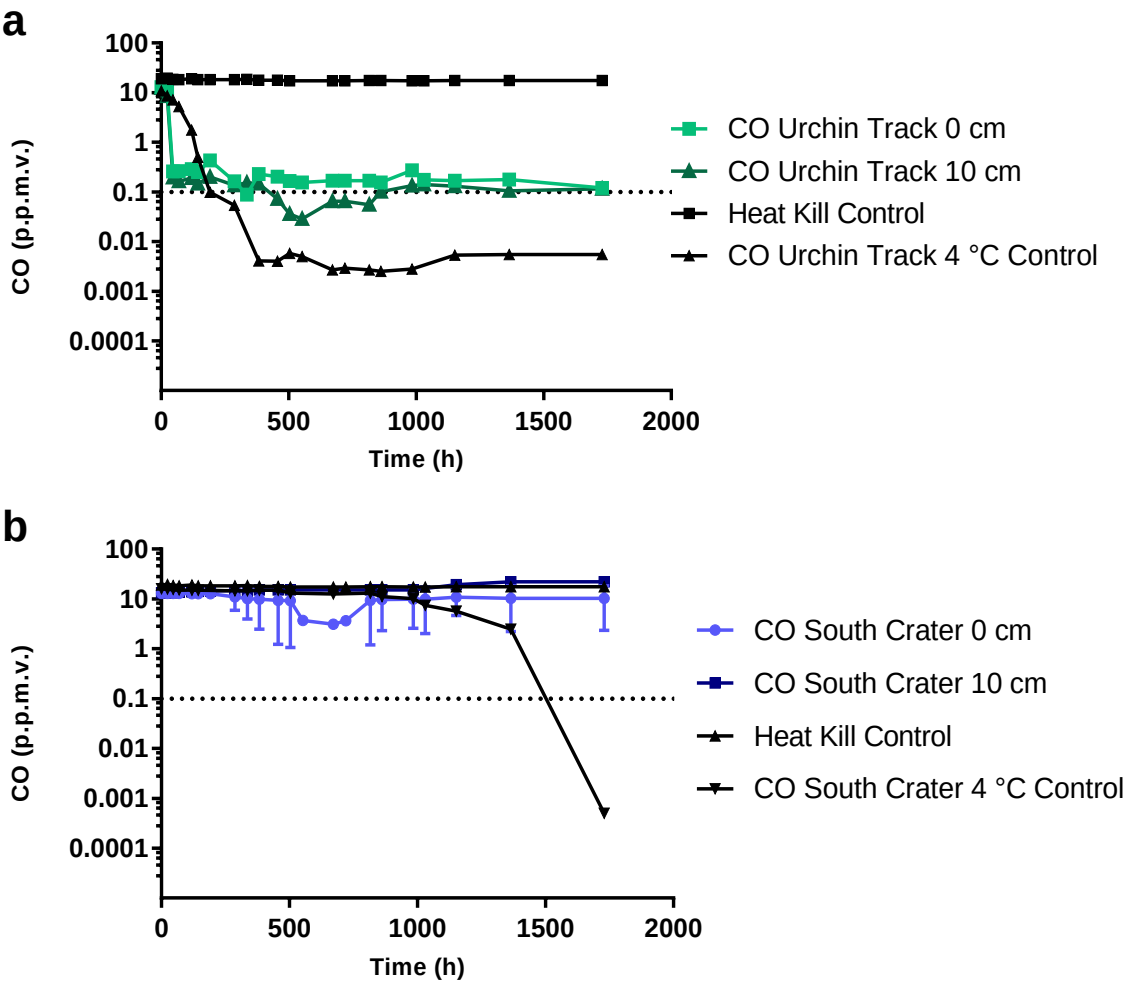

Figure 27: Thermophilic trace gas oxidation of CO in volcanic (Mt. Tongariro) and volcanic-adjacent (Mt. Urchin) soil microcosms. Gas chromatography measurements of 60 °C incubations tracking oxidation of atmospheric CO (mixing ratio, 0.1 parts per million by volume (p.p.m.v.) at **a,** Mt. Urchin Track, Kaimanawa Forest Park and **b,** South Crater, Mt. Tongariro.

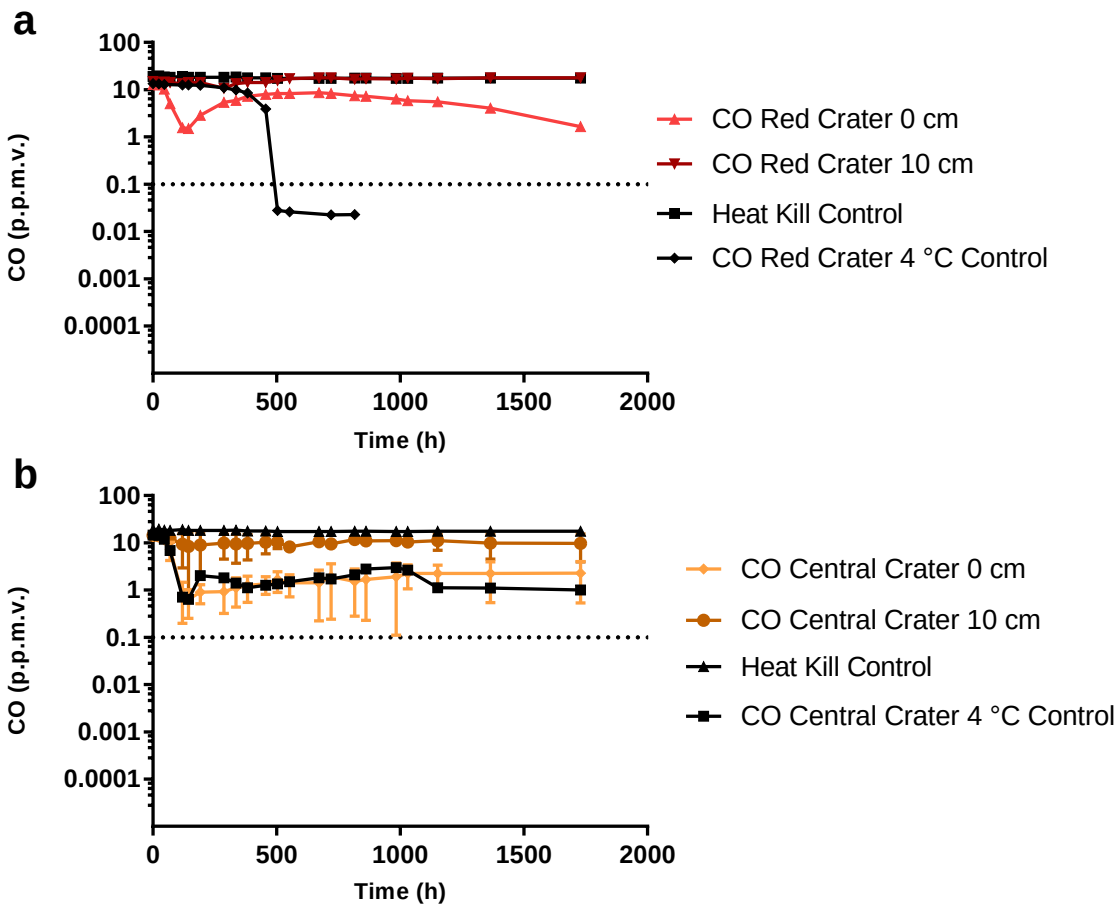

Figure 28: Thermophilic trace gas oxidation of CO in volcanic (Mt. Tongariro) soil microcosms. Gas chromatography measurements of 60 °C incubations tracking oxidation of atmospheric CO (mixing ratio, 0.1 parts per million by volume (p.p.m.v.) at **a,** Red Crater, Mt. Tongariro and **b,** Central Crater, Mt. Tongariro.

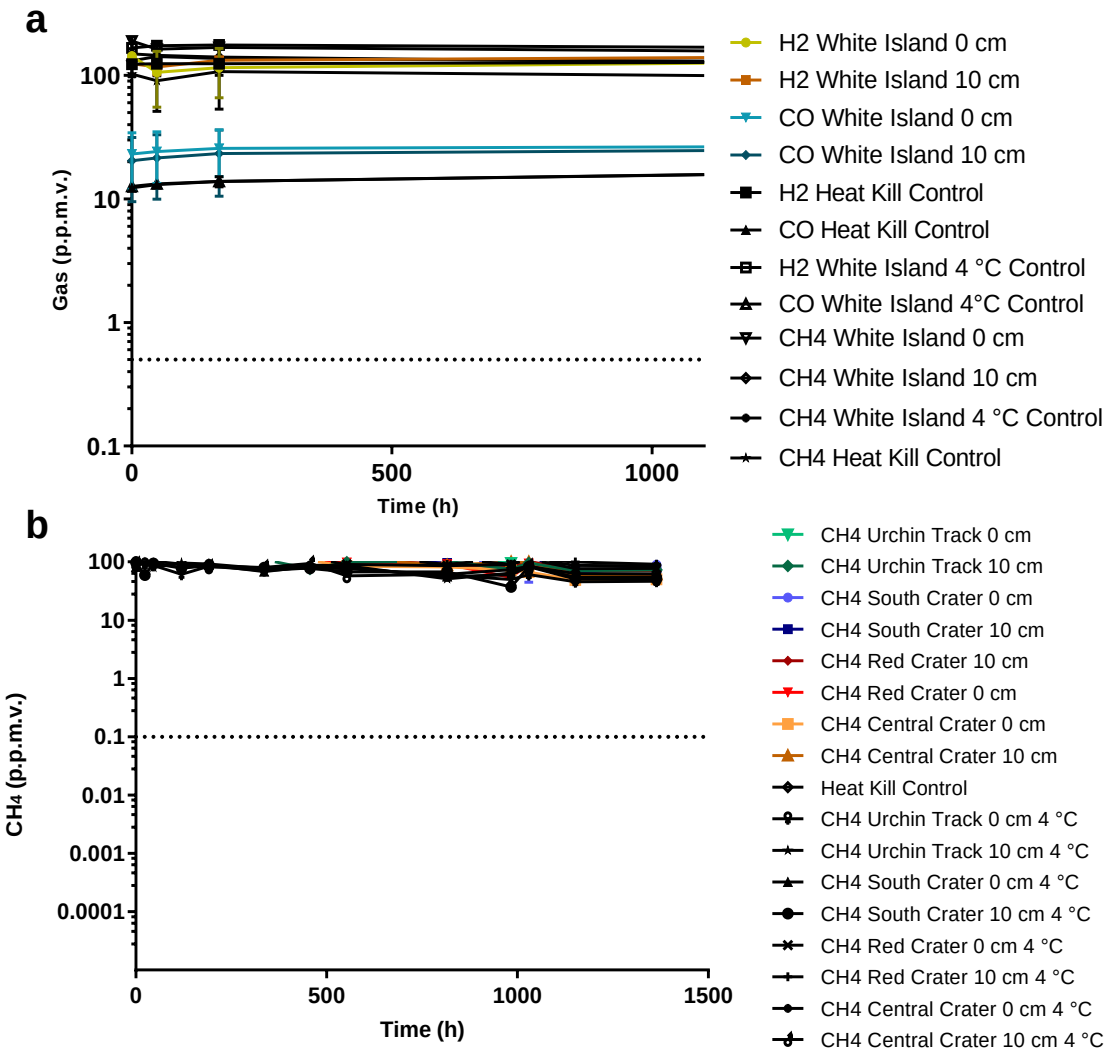

Figure 29: Thermophilic trace gas oxidation of CO, H_2 and CH_4 in volcanic (White Island) samples and CH_4 oxidation in Mt. Urchin, Mt. Tongariro. Gas chromatography measurements of 60 °C incubations tracking oxidation of atmospheric CO (mixing ratio, 0.1 p.p.m.v.) H_2 (mixing ratio 0.5 p.p.m.v.) and CH_4 (mixing ratio, 1.8 p.p.m.v.) at **a,** White Island and **b,** Mt. Urchin, Kaimanawa Forest Park and Red South and Central Craters, Mt. Tongariro at 60 °C.

3.4 Discussion: Microbial Trace Gas Oxidation

3.4.1 Discussion: Trace Gas Oxidation at 4 °C

Oxidation of tropospheric trace gases H_2 and CO are dependable sources of energy for taxa in oligotrophic environments such as the Antarctic Dry Valleys, Atacama Desert, and bare volcanic rocks in Iceland and Hawaii (see Section 1.1.4 and 1.3)[1,3,37,184]. Taxa with metabolisms ranging from chemolithoautotrophy, mixotrophy to heterotrophy in starvation are capable of utilising H_2 and CO[1,44,45]. Whether trace gas oxidation is a common mechanism of persistence for all terrestrial communities, including volcanic and oligotrophic systems devoid of vascular plants is an interesting question.

Evaluating the results from our trace-gas oxidation experiments using Mt. Tongariro and bare soils controls, it is likely that aerobic CO and H_2 oxidation are reliable sources of energy for persisting taxa in bare soil habitats. Furthermore, the presence of carbon in soils is potentially influencing the uptake of CO and H_2, consistent with literature. Literature shows a correlation between amount of organic carbon and greater abundance and diversity of CO oxidizers. In addition, more rapid rates of CO oxidation are associated with carbon in the system[185]. Also, consistent with the results in the literature, no soils oxidised CH_4, despite the methanotrophic genus *Methylacidiphilum* being one of the most abundant genera in detected in the soil survey dataset (mean abundance of 53.9%; Supplemental Table 4) and the most prevalent genera (found in all 30 soil samples). This lack of CH_4 oxidation was also observed in microbial communities from bare volcanic cinders from Hawaii[3]. On the other hand, Mt. Tongariro and non-volcanic bare soil controls were capable of oxidising H_2 and CO to under atmospheric levels (< 0.5 p.p.m.v. and < 0.1 p.p.m.v., respectively). This is consistent with other studies of volcanic areas[186]. Notably, in soils from Mt. Urchin, Red Crater, Central Crater and Mt. Sugarloaf, complete oxidation of H_2 was more rapid at the surface (0 cm) compared to depth (10 cm).

Oxidation occurring more rapidly at the surface is logical from the following perspective: soils at surface are more aerated than soils at depth[187] meaning they likely harbour more aerobic taxa, and aerobic taxa contain the high-affinity respiratory linked [NiFe]-hydrogenase enzymes required for trace gas oxidation.

These [NiFe]-hydrogenase enzymes are oxygen-tolerant lineages (groups 1d, 1h and 2a) and are commonly detected in aerated samples such as forest and agricultural soils[66]. This suggests that there could be a link between exposure to the atmosphere and prevalence of oxygen-tolerant hydrogenase enzymes, therefore, it is more likely that trace-gas oxidation is a primed capability for surface soil communities where these is exposure to atmospheric H_2 and CO, particularly in soils where active fumarolic degassing is absent[2,66]. Mt. Tongariro and the control site soils all returned type 1h [NiFe] hydrogenase sequences with the most similarity to *Streptomyces* sequences in the phylum Actinobacteria (Figure 22). Many *Streptomyces* taxa are associated with encoding Type 1h [NiFe] hydrogenases[40,114]. Other Actinobacteria lineages from the genus *Pseudonocardia* are known trace gas (CO and H_2) oxidisers. The genus *Pseudonocardia* is present in 26/30 soil samples and has a mean abundance at genera level of 28%. *Pseudonocardia* encode Group 1h [NiFe] hydrogenase[178,188] and *Pseudonocardia* 16S rRNA gene sequences[4] have been detected in other high elevation samples from Antarctic mineral soils, Atacama Desert soils and Himalayan high-elevation soils[158] and from Icelandic volcanic deposits[40].

High affinity aerobic type I [MoCu] carbon monoxide dehydrogenase enzyme genes were detected in Mt. Tongariro and control soils, returning sequence similarity to novel, uncultivated taxa (Figure 22). These CO-oxidising enzymes are associated with carboxydovores, capable of oxidising low levels of CO (1% v/v or 10,000 p.p.m.v.). In contrast, carboxydotrophs encode low affinity nickel enzymes which are only active in elevated (>1% v/v or 10,000 p.p.m.v.) CO concentrations[46]. Evaluating these results, it is likely that aerobic CO oxidisers are responsible for atmospheric CO oxidation in these samples and may be using CO as a supplemental energy source in times of carbon limitation. Interestingly, soils from Central Crater, Red Crater, Mt. Urchin and Mt. Sugarloaf all contain detectable levels of total carbon ranging from 600-16,900 mg/kg (Supplementary Table 8) and soils from these sites also oxidised CO more rapidly. This corroborates studies where increased organic carbon concentrations are correlated with greater abundance and diversity of CO oxidisers[185]. Therefore, organic carbon availability may be responsible for the higher CO uptake rates. Also, most described CO oxidizers are heterotrophs, which use CO as a source of supplementary energy in C-limitation.

In carbon limitation, CO becomes a dependable source of energy for cell maintenance[189].

3.4.2 Discussion: Trace Gas Oxidation at 60 °C

One of the interesting observations made as part of the soil community surveys was the detection of putatively thermophilic taxa within these cold ecosystems. To determine whether these taxa were thermophilic and/or able to respire trace gases at their reported optimal temperatures, the same soil microcosms used for 4 °C experiments were incubated at 60 °C and the concentrations of H_2, CO and CH_4 were tracked. One of the technical or biological triplicates was tracked at 4 °C to as a control, to ensure oxidation (or lack thereof) was the consequence of biological processes, not due to a lack of oxygen for aerobic oxidation of H_2 or CO. No CH_4 oxidation was detected at any site at 60 °C. Previous molecular and ecological studies show that cinder CO oxidising communities consist predominantly of taxa closely related to class *Ktedonobacteria*[185]. Chloroflexi from the class include thermophilic or thermotolerant strains, such as *Thermogemmatispora carboxydovorans*, and *Thermogemmatispora* sp. T81[157,190]. Interestingly, a previous study reports that CO oxidation does not occur in bare volcanic soils above 55 °C yet it does occur in forest canopy soils. Conversely, H_2 oxidation occurred in bare soils at 55 °C[186]. In line with this, soil microcosms from bare Mt. Urchin soil were capable of complete H_2 oxidation to under 0.5 p.p.m.v. (Figure 22, graph a) in under 500 hours. However, for Mt. Tongariro, only soils at 10 cm depth for South Crater were capable of oxidising H_2 to under 0.5 p.p.m.v. *in situ*. South Crater samples (at the surface) were capable of oxidising H_2 from 100 p.p.m.v. to ~10 p.p.m.v. but were not capable of high-affinity oxidation, indicating perhaps a lack of thermophilic taxa which encode high-affinity [NiFe]-hydrogenases necessary for respiration of atmospheric H_2. These results indicate hydrogenase activity is still present, but perhaps the thermophilic taxa encode low affinity hydrogenases which are incapable of oxidising high levels of H_2 and are lack the affinity to H_2 for complete atmospheric oxidation. This could be because the soils from Mt. Tongariro which could be due to the previous history of the South, Red and Central Craters which previously active degassing of fumaroles would have been present in these sites[6].

Similarly, complete oxidation of CO at 60 °C was only observed for Mt. Urchin (Figure 24, a). Mt. Tongariro soils at Red Crater (0 cm) and Central Crater (0 cm) (Figure 25) were both capable of oxidising CO to around 1 p.p.m.v., but not to under atmospheric concentrations (<0.1 p.p.m.v.). These results indicate either that the enzymes active at thermophilic temperatures are low-affinity enzymes, and that the taxa responsible for complete oxidation of H_2 and CO in the 4 °C experiments are potentially different taxa than those active at thermophilic (60 °C) temperatures; or perhaps the levels of carbon were depleted in the soils, and therefore mixotrophy by the carboxydovores in the microcosms was reduced, as CO-oxidation is often a source of supplemental energy and other sources of carbon are necessary for biomass[186]. However, previous studies of volcanic areas have returned CO-oxidising potential thermophiles, but have failed to demonstrate active CO-oxidation above 55 °C in bare, volcanic soils[186,190].

3.5 Conclusions: Trace Gas Oxidation at 4 °C and 60 °C

Volcanic soils of Mt. Tongariro and non-volcanic bare soils of control sites Mt. Sugarloaf and Mt. Urchin were all capable of oxidising H_2 and CO. At 4 °C soil microcosms were shown to all oxidise H_2 and CO to below atmospheric trace levels (0.5 and 0.1 p.p.m.v., respectively). In addition, gene sequences were able to be extracted from the communities using specific primers for Type 1h [NiFe] dehydrogenase and Type 1 [MoCu] carbon monoxide dehydrogenase large subunit primers, showing their presence.

Conversely, at 60 °C the only site which oxidised both H_2 and CO to under atmospheric levels were soils from Mt. Urchin. South Crater, Mt. Tongariro at 10 cm depth also oxidised H_2 to under atmospheric levels yet all other sites were not capable of full oxidation of H_2 or CO. Sites with partial H_2 or CO oxidation are likely to have low affinity enzymes which are suited to elevated concentrations of H_2 and CO; which would make these taxa better suited to areas where volcanic degassing and thermophily are occurring: for Mt. Tongariro, where in the past, active degassing released greater concentrations of H_2 and CO into the environment, therefore, the thermophilic taxa may never have adapted to have high affinity enzymes.

Chapter 4: Summary Section

4.1 Intro

The crux of this thesis was to address the microbial ecology and persistence mechanisms of soil taxa in non-active volcanic soils from Mt. Tongariro. It was hypothesised that soil taxa from Mt. Tongariro and other bare soil types would oxidise atmospheric trace gases (H_2, CO) to persist in times of carbon limitation and that thermophilic taxa were present in the soils and persisting using H_2 and CO as sources of energy in dormancy. To demonstrate this hypothesis, three research objectives were accomplished. The first was to detail the microbial ecology and physicochemical status of bare volcanic and non-volcanic control soils to understand the community composition and likely inferred function of said communities. Secondly, trace gas oxidation at 4 °C was detailed to explore mechanisms of persistence in these extreme volcanic soils. Lastly, to address whether thermophiles persist by oxidising atmospheric trace gases in Mt. Tongariro soils, trace-gas oxidation at 60 °C was accomplished.

4.2 Research Significance

This thesis is the first investigation into the microbial ecology in Tongariro National Park. It gives preliminary insights into the composition and ecological function of microbial communities residing in bare, oligotrophic volcanic soils of Mt. Tongariro. This research indicates that some thermophilic taxa are still active on Mt. Tongariro and that thermophilic taxa can oxidise trace gas both in some volcanic and temperate soil types. This research corroborates studies from Antarctic Dry Valleys and metagenomic inferences and other microbial ecology studies from the Atacama Desert which postulate that trace-gas oxidation is a source of energy for microbial taxa persisting in suboptimal conditions. In arid systems, the mode of primary production in lieu of phototrophy: either prokaryotic or eukaryotic is postulated to be trace-gas oxidation[1,2,37]. However, the presence of moisture and Cyanobacteria in Mt. Tongariro soils means trace-gas oxidation is likely a source of supplemental energy to mixotrophic taxa and Cyanobacteria are the primary producers in the system. This could explain why the observed diversity

and microbial abundance of Mt. Tongariro and the control soils is comparable to other temperate soil environments.

4.3 Future Directions

There are several limitations with this research which are due to time limitation and scope. Firstly, the amount of sample obtained for analyses per site from Mt. Tongariro was approximately 20 g. This meant that soil had to be recycled for repeat trace gas oxidation experiments. If unconstrained by time and money, metagenomic and metatranscriptomic analyses of samples and qPCR of RNA extracted from Mt. Tongariro soils once enriched with trace gases would have been excellent to see the activity of specific enzyme uptake genes in the soils. It would have been more informative to the study to have actively transcribed genes, not just presence or absence of the function.

Additionally, the extraction method, the sequencing technology, sequencing chemistry used, and downstream analysis of sequencing outputs will always have biases, no matter the methods used. For instance, 16s community analysis using Illumina Miseq technology which sequences only the V4 region of the 16s rRNA gene, a short 250 bp length fragment. This undoubtably makes for taxonomic assignment error as the length of sequence is very small. Using the V4 region, species assignment is unreliable. The most straightforward way to look at V4 sequence data is to only consider the phylum or genus level accurate.

For future studies, it would be wise to consider the addition of metagenomes so that the function of a community could be assessed, instead of relying on inferred function (from bioinformatics software such as PICRUSt) from known isolates using 16S rRNA genes. Analysis of metagenomes would give a more accurate picture of the functional capability of Mt. Tongariro soils. Also, the addition of a screening test for relic DNA would be necessary to rule out bare DNA which has adsorbed to soil particles and obscuring observed microbial diversity[191]. Thermophilic isolates have been shown to oxidise H_2 and CO in times of carbon limitation at 60 °C[44,45], though, it would be interesting to see if thermophilic isolates could oxidise H_2 and CO at sub-optimal temperatures such as 4 °C. Similarly, to give more insight into the functional capability of the environment, screening for RuBisCO genes which are present in carboxydotrophic taxa and CO_2 oxidising

bacteria would be excellent to have a better picture of the ecology of the trace gas oxidising taxa present[1]. Screening for low affinity hydrogenase and carbon monoxide dehydrogenase enzymes would be necessary to give a more sophisticated analysis of the trace gas ecology and how microbial populations are persisting in these bare volcanic soils.

Additionally, experiments in which isolates, and soil microcosms are enriched with carbon and oxidation of H_2 and CO are tracked, in tandem to isolates and soil microcosms (without carbon enrichment) are necessary to see whether the taxa oxidising trace gas are true scavengers, persisting in a state of dormancy by means of trace gas oxidation or whether they are mixotrophic taxa concurrently oxidising trace gases and reducing organic carbon in the system and are active and replicating.

4.4 Final Conclusions

Soils of Mt. Tongariro harbour surprising biodiversity, likely due to the increased moisture input and presence of Cyanobacteria compared to other oligotrophic volcanic and hyperarid systems. Also, certain physicochemical variables, such as altitude, silica, sodium and lithium appear to be decent predictors of microbial diversity across the dataset, however, only 28% of the diversity can be correlated to physicochemistry. H_2 and CO seem to be dependable sources of energy for taxa persisting on the Tongariro Alpine Crossing, yet, it is inconclusive whether atmospherically relevant trace-gas oxidation at thermophilic temperatures 60 °C is occurring on the non-active South, Red and Central Craters of Mt. Tongariro. It is possible that the thermophilic taxa are capable to oxidising elevated levels of H_2 and CO but lack high affinity enzymes required for complete oxidation of atmospherically relevant concentrations of trace gases.

In conclusion, this study is the first preliminary study of Mt. Tongariro. The microbial composition of these volcanic soils are comparable to other volcanic soils which have detectable soil moisture[2], yet differ from hyperarid oligotrophic systems due to the presence of Cyanobacteria[37,145]. Overall, this thesis achieved its purpose: to detail the microbial diversity and the utilisation of trace-gas oxidation as a persistence mechanism. Like all preliminary studies, there is much work left to do for this story to be robust and told in higher resolution.

SUPPLEMENTARY TABLES

Supplementary Table 1 | New Zealand Geothermal Analytical Laboratory (NZGAL) detection method and detection limit for each chemical species.

Chemical Species	Method	Detection Limit	
Ammonium	FIA – APHA 4500 NH3-H 22nd Edition 2012	0.003	mg/kg
Boron	ICP-OES, GW, APHA 3120-B 22nd Edition 2012	5	mg/kg
Caesium	ICP-MS, Trace Level, APHA 3125-B 22nd Edition 2012	0.005	mg/kg
Calcium	ICP-OES, GW, APHA 3120-B 22nd Edition 2012	1	mg/kg
Chloride	Ion Chromatography – APHA 4110-B 22nd Edition 2012	0.05	mg/kg
Cobalt	ICP-MS, Trace Level, APHA 3125-B 22nd Edition 2012	0.005	mg/kg
Copper	ICP-MS, Trace Level, APHA 3125-B 22nd Edition 2012	0.005	mg/kg
Iron	ICP-OES, GEO, APHA 3125-B 22nd Edition 2012	2	mg/kg
Lithium	ICP-MS, Trace Level, APHA 3125-B 22nd Edition 2012	0.05	mg/kg
Magnesium	ICP-OES, GEO, APHA 3120-B 22nd Edition 2012	1	mg/kg
Manganese	ICP-MS, Trace Level, APHA 3125-B 22nd Edition 2012	0.05	mg/kg
Molybdenum	ICP-MS, Trace Level, APHA 3125-B 22nd Edition 2012	0.05	mg/kg
Nickel	ICP-MS, Trace Level, APHA 3125-B 22nd Edition 2012	0.005	mg/kg
Nitrogen (as N)	Ion Chromatography – APHA 4110-B 22nd Edition 2012	0.01	mg/kg
Phosphorus	FIA, APHA 4500-PG (Modified) 22nd Edition 2012	0.002	mg/kg
Potassium	ICP-OES, GW, APHA 3120-B 22nd Edition 2012	10	mg/kg
Total S (as SO_4)	ICP-OES - APHA 3120-B 22nd Edition 2012	45	mg/kg
Silica	ICP-OES, GW, APHA 3120-B 22nd Edition 2012	10	mg/kg
Sodium	ICP-OES, GW, APHA 3120-B 22nd Edition 2012	2	mg/kg
Sulphate	Ion Chromatography – APHA 4110-B 22nd Edition 2012	0.03	mg/kg
Zinc	ICP-MS, Trace Level, APHA 3125-B 22nd Edition 2012	0.05	mg/kg

Supplementary Table 2 | All physicochemical values, presented in mg/k. NA = not applicable.

Chemical Species	Site ID	Urchin		South Crater		Red Crater		Central Crater		White Island		Mt Sugarloaf 1		Mt Sugarloaf 1	
	Sample ID	EQC1-0	EQC1-10	EQC2-0	EQC2-10	EQC5-0	EQC5-10	EQC8-0	EQC8-10	EQC13-0	EQC13-10	EQC_11_0	EQC11_10	EQC12_0	EQC12_10
Ammonium as N	mg/kg	0.5	<0.1	0.3	<0.1	<0.1	<0.1	0.2	<0.1	0.1	<0.1	0.466	0.291	0.994	NA
Boron	mg/kg	<5.0	<5.0	<5.0	<5.0	<5.0	<5.0	<5.0	<5.0	<5.0	<5.0	<40	<40	<40	<40
Caesium	mg/kg	1	0.87	0.45	0.36	0.18	0.12	0.23	0.24	0.03	0.37	NA	NA	NA	NA
Calcium	mg/kg	2472	3453	3535	3365	5087	4893	7910	5481	13823	3802	5565.00	3821.00	7080.00	6895.00
Chloride	mg/kg	<5	<5	5	<5	<5	<5	6	10	47	148	<5	<5	<5	<5
Cobalt (Trace Level)	mg/kg	2.1	2.1	2.4	2.3	5.5	10.3	4.1	3.8	3.1	0.39	12.2	8.3	10.9	11.5
Copper (Trace Level)	mg/kg	12.8	15	16.3	16.9	19.8	22	19	14	12.3	7.5	25.0	20.0	15.6	14.3
Diss. Reactive Phosphorus	mg/kg	4	4	5.3	4.2	52	46	16.3	19.1	<1	<1	74.0	29.0	169.0	NA
Iron	mg/kg	16438	16792	22274	24303	8770	10630	16686	12405	7728	12324	35087	39152	29915	29716
Lithium (Trace Level)	mg/kg	3.3	4	1.3	1.1	0.78	0.8	1.8	1.5	0.08	0.9	59.0	49.0	44.0	38.0
Magnesium	mg/kg	497	581	1293	1153	4868	9688	2777	3352	19.2	301	8395	7286	8049	8492
Manganese (Trace Level)	mg/kg	120	117	45	39	82	117	67	65	0.8	5.3	566	422	429	409
Molybdenum (Trace Level)	mg/kg	0.35	0.36	0.2	0.23	<0.05	0.05	0.18	0.14	0.28	0.21	0.14	0.10	0.15	0.12
Nickel (Trace Level)	mg/kg	2.4	2.4	4.7	3.8	13.4	31	7.7	8.7	8.6	1.3	24.0	26.0	20.0	40.0
Nitrate as N	mg/kg	<1	<1	<1	<1	<1	<1	<1	<1	<1	<1	2.0	<1	<1	<1
Potassium	mg/kg	233	234	365	253	140	196	331	186	261	1927	1642	1197	1082	1025
Silica (as SiO2)	mg/kg	9679	8960	3033	2199	1747	2588	3377	2272	1352	1559	367	2775	225	151
Sodium	mg/kg	611	898	847	742	883	1034	1585	1106	122	777	187	168	192	359
Sulphate	mg/kg	22	61	37	68	21	33	54	62	4939	13287	13	6	8	10
Total S (as SO4)	mg/kg	1531	1910	2875	2323	381	212	2110	1141	83296	21740	239	193	153	152
Zinc (Trace Level)	mg/kg	14.3	13.8	8.5	7.7	9.1	10.3	11.3	9.1	0.43	1.1	116	101	87	85
TIC	mg/kg	<800	<800	900	<800	<800	<800	<800	<800	<800	<800	<1800	<1800	<1800	<1800
TC	mg/kg	13700	16900	1300	<500	600	<500	1100	1000	<500	<500	3700	6700	1500	<1300
TOC	mg/kg	13600	18200	<500	<500	<500	<500	800	900	<500	<500	5400	6500	1600	1300

Supplementary Table 3 | Shows phyla with more than an average threshold of 1%. Maximum and mean abundance of each phyla and prevalence per sample and standard deviation across all 30 samples is listed.

phylum	mean abundance	SD	max	prevalence
Proteobacteria	0.202	0.107	0.408	30
Chloroflexi	0.182	0.107	0.386	30
Verrucomicrobia	0.105	0.106	0.436	30
Acidobacteria	0.078	0.037	0.135	30
Thaumarchaeota	0.062	0.026	0.112	30
Actinobacteria	0.051	0.044	0.129	30
Planctomycetes	0.048	0.028	0.147	30
Crenarchaeota	0.046	0.039	0.116	30
Deinococcus-Thermus	0.034	0.035	0.160	30
WPS-2	0.028	0.024	0.072	30
Cyanobacteria	0.026	0.043	0.242	30
Armatimonadetes	0.023	0.016	0.074	30
Elusimicrobia	0.020	0.022	0.087	30
Epsilonbacteraeota	0.018	0.019	0.083	29
Bacteroidetes	0.011	0.007	0.025	30
Gemmatimonadetes	0.011	0.006	0.033	30
GAL15	0.011	0.016	0.067	30

Supplementary Table 4 | Top genera with more than 1% average threshold are shown. Maximum and mean abundance calculated at the genera level. Prevalence per sample and standard deviation across all 30 samples is listed for reads at the genera taxonomic level. If the genera are unclassified, the genus is displayed by the next available classified taxonomic rank: f = family, o = order, p = phylum.

phylum	*genus*	mean abundance	SD	max	prevalence
Verrucomicrobia	*Candidatus Methylacidiphilum*	0.054	0.073	0.342	30
Verrucomicrobia	*Candidatus Udaeobacter*	0.052	0.117	0.472	29
Deinococcus-Thermus	*Thermus*	0.030	0.037	0.171	30
Actinobacteria	*Pseudonocardia*	0.028	0.037	0.094	26
Chloroflexi	*AD3 (c)*	0.026	0.032	0.098	30
Crenarchaeota	*Geoarchaeales (o)*	0.023	0.020	0.062	30
Elusimicrobia	*Elusimicrobia (c)*	0.023	0.026	0.101	29
Chloroflexi	*Anaerolineaceae (f)*	0.023	0.028	0.090	30
Chloroflexi	*AD3 (o)*	0.022	0.024	0.079	29
Planctomycetes	*Gemmataceae (f)*	0.022	0.018	0.104	30
Chloroflexi	*SBR1031 (o)*	0.020	0.027	0.069	23
Proteobacteria	*Acidithiobacillus*	0.016	0.017	0.073	28
Crenarchaeota	*Sulfolobaceae (f)*	0.016	0.016	0.045	27
Thaumarchaeota	*Nitrosopumilaceae (f)*	0.015	0.020	0.050	18
WPS-2	*WPS-2 (p)*	0.015	0.017	0.052	30
Armatimonadetes	*Chthonomonas*	0.014	0.017	0.075	30
Proteobacteria	*Gallionella*	0.014	0.046	0.244	20
Proteobacteria	*Thiobacillus*	0.014	0.014	0.058	28
Actinobacteria	*Crossiella*	0.014	0.017	0.053	26
Acidobacteria	*Blastocatellia (Subgroup 4) (c)*	0.013	0.017	0.042	19
WPS-2	*WPS-2 (p)*	0.012	0.016	0.064	27
Chloroflexi	*P2-11E (c)*	0.012	0.030	0.165	29
Proteobacteria	*Chromobacterium*	0.012	0.014	0.061	30
Thaumarchaeota	*Candidatus Nitrosotalea*	0.012	0.017	0.079	28
Epsilonbacteraeota	*Sulfurimonas*	0.012	0.014	0.062	28
Thaumarchaeota	*Candidatus Nitrocosmicus*	0.011	0.018	0.068	25
Gemmatimonadetes	*Gemmatimonadaceae (f)*	0.010	0.007	0.038	30

Supplementary Table 5 | Linear regression of physicochemical parameters individually against alpha diversity (Shannon diversity index). List is ordered by lowest to highest value.

	Slope	Std. Error	t-value	-value	Signif. Codes	R^2
Silica	1.236e-04	3.817e-05	3.238	0.0031	*	0.2724
Caesium	0.8778	0.3246	2.704	0.0124	*	0.2335
TC	0.0000	0.0000	2.166	0.0390	*	0.1435
Calcium	-0.0001	0.0000	-2.083	0.0465	*	0.1341
Nickel	-0.0162	0.0080	-2.029	0.0520	.	0.1282
TOC	0.0000	0.0000	1.932	0.0636	.	0.1176
Altitude	0.0002	0.0001	1.868	0.0722	.	0.1108
Sodium	0.0004	0.0002	1.859	0.0736	.	0.1098
Sulfur	0.0000	0.0000	-1.624	0.1160		0.0861
Soil Moisture	0.0136	0.0105	1.296	0.2050		0.0566
Cobalt	-0.0316	0.0250	-1.260	0.2180		0.0536
Lithium	-0.0067	0.0055	-1.227	0.2300		0.0510
Copper	0.0248	0.0202	1.228	0.2300		0.0511
Zinc	-0.0033	0.0028	-1.170	0.2520		0.0466
Potassium	-0.0002	0.0002	-1.104	0.2790		0.0417
Temp	-0.0062	0.0066	-0.949	0.3510		0.0312
Sulfate	0.0000	0.0000	-0.941	0.3550		0.0306
Manganese	-0.0006	0.0006	-0.937	0.3570		0.0304
Chloride	-0.0018	0.0020	-0.886	0.3830		0.0273
Magnesium	0.0000	0.0000	-0.875	0.3890		0.0266
Molybdenum	0.7935	0.9426	0.842	0.4070		0.0247
Nitrate	0.1394	0.2483	0.561	0.5790		0.0111
pH	0.0389	0.0762	0.511	0.6130		0.0093
Phosphorus	-0.0009	0.002191	-0.418	0.6790		0.0064
TIC	-0.0001311	0.0003	-0.396	0.6950		0.0056
Iron	-1.441e-06	0.0000	-0.135	0.8940		0.0006
Ammonium	-0.0183	0.3437	-0.053	0.9580		0.0001

Signif. codes: 0-0.001 '***'; 0.001-0.01 '**'; 0.01-0.05 '*'; 0.05-0.1 '.'; 0.1-1 ' '

Supplementary Table 6 | Multiple linear regression model of significant physicochemical parameters against alpha diversity (Shannon diversity index). Collinear variables were removed and an Akaike information criterion (AIC) was applied to the variables. Variables were selected for the model in order of best fit from highest to lowest from singular linear regression calculations.

	Estimate	**Std. Error**	**t-value**	**-value**	**Signif. codes**
(Intercept)	1.205e+01	3.090e+00	3.899	0.00105	**
Silica	1.521e-03	6.979e-04	2.180	0.04279	*
Calcium	-5.060e-04	2.34E-04	-2.159	0.0446	*
Nickel	-1.23E-01	5.815e-02	-2.113	0.04882	*
Caesium	-1.53E+01	7.35E+00	-2.084	0.05166	.
Sodium	-1.16E-03	6.32E-04	-1.842	0.08208	.
TOC	-1.325e-04	8.262e-05	-1.604	0.12621	
Altitude	-4.13E-04	2.95E-04	-1.399	0.17884	

Multiple R-squared: 0.487, adjusted R-squared: 0.2875

F-statistic: 2.441 on 7 and 18 degrees of freedom, p-value: 0.06023

Signif. codes: 0-0.001 '***'; 0.001-0.01 '**'; 0.01-0.05 '*'; 0.05-0.1 '.'; 0.1-1 ' '

Supplementary Table 7 | Mantel tests using Spearman's correlation (permutations = 999) of Bray-Curtis distance matrix (dissimilarities) between all communities sampled and each individual physicochemical parameter and altitude.

	Rho (ρ)	**-value**	**Signif. codes**
Lithium	0.5631	0.001	***
TOC	0.5223	0.001	***
Zinc	0.4988	0.001	***
Manganese	0.4876	0.001	***
Iron	0.3712	0.001	***
Altitude	0.2962	0.001	***
pH	0.2622	0.001	***
Sodium	0.2532	0.001	***
Ammonium	0.3159	0.002	**
Phosphorus	0.3125	0.002	**
Magnesium	0.2671	0.002	**
Air Temp	0.2496	0.002	**
TC	0.2804	0.003	**

Cobalt	0.2777	0.005	**
Potassium	0.2715	0.005	**
Silica	0.2371	0.005	**
Nickel	0.234	0.007	**
Temp	0.2045	0.008	**
Sulfate	0.2085	0.011	*
Calcium	0.1686	0.021	*
Copper	0.1545	0.025	*
Sulfur	0.1547	0.032	*
Nitrate	0.2472	0.039	*
Chloride	0.09008	0.12	
Soil Moisture	0.07456	0.145	
Caesium	0.083	0.159	
Molybdenum	0.054	0.217	
TIC	-0.1132	0.942	

Signif. codes: 0-0.001 '***'; 0.001-0.01 '**'; 0.01-0.05 '*'; 0.05-0.1 '.'; 0.1-1 ' '

Supplementary Table 8 | Calculations for Total Organic Carbon (TOC), Total Inorganic Carbon (TIC) and Total Carbon (TC) for all sample sites.

sample site	site ID	TIC mg/kg	TC mg/kg	TOC mg/kg
Urchin	EQC1-0cm	<800	13700	13600
	EQC1-10cm	<800	16900	18200
South Crater	EQC2-0cm	900	1300	<500
	EQC2-10cm	<800	<500	<500
Red Crater	EQC5-0cm	<800	600	<500
	EQC5-10cm	<800	<500	<500
Central Crater	EQC8-0cm	<800	1100	800
	EQC8-10cm	<800	1000	900
White Island	EQC13-0cm	<800	<500	<500
	EQC13-10cm	<800	<500	<500
Mt Sugarloaf 1	CASS1_0cm	<1800	3700	5400
	CASS1_10cm	<1800	6700.00	6500.00
Mt Sugarloaf 2	CASS2_0cm	<1800	1500.00	1600.00
	CASS2_10cm	<1800	<1300	1300.00

Supplementary Table 9 | Calculations for soil moisture, as outlined in methods (Section 2.2.2).

Site	Sample	Tray	Soil + Tray	Wet Soil Weight	Dry Soil + Tray	Dry Soil Weight	% Moisture
Urchin	EQC_1_0cm	0.9805	2.0059	1.0254	1.7268	0.7463	27.2
	EQC_1_10cm	0.942	2.1775	1.2355	1.565	0.623	49.6
South Crater	EQC_2_0cm	1.0809	2.0434	0.9625	1.8327	0.7518	21.9
	EQC_2_10cm	0.9836	2.0757	1.0921	1.9182	0.9346	14.4
	EQC_3_0cm	1.0531	2.2239	1.1708	1.8143	0.7612	35.0
	EQC_3_10cm	1.0139	2.0671	1.0532	1.8714	0.8575	18.6
	EQC_4_0cm	1.0595	2.0095	0.95	1.865	0.8055	15.2
	EQC_4_10cm	1.0782	2.2894	1.2112	2.0307	0.9525	21.4
	EQC_5_0cm	1.0391	2.1463	1.1072	2.0251	0.986	10.9
	EQC_5_10cm	1.0074	2.3806	1.3732	2.1525	1.1451	16.6
Red Crater	EQC_6_0cm	1.0396	2.1441	1.1045	2.0521	1.0125	8.3
	EQC_6_10cm	1.0023	2.0479	1.0456	1.9879	0.9856	5.7
	EQC_7_0cm	1.0182	2.3344	1.3162	2.1813	1.1631	11.6
	EQC_7_10cm	1.0603	2.0365	0.9762	1.9433	0.883	9.5
	EQC_8_0cm	1.0194	2.8447	1.8253	2.5439	1.5245	16.5
	EQC_8_10cm	0.9802	2.0033	1.0231	1.834	0.8538	16.5
Central Crater	EQC_9_0cm	1.013	2.1772	1.1642	1.9751	0.9621	17.4
	EQC_9_10cm	1.021	2.3333	1.3123	2.0656	1.0446	20.4
	EQC_10_0cm	0.9941	2.2428	1.2487	2.1036	1.1095	11.1
	EQC_10_10cm	1.0372	2.0683	1.0311	1.9065	0.8693	15.7
	EQC_13_0cm	1.1309	2.1887	1.0578	1.962	0.8311	21.4
	EQC_13_10cm	1.1213	2.303	1.1817	2.232	1.1107	6.0
White Island	EQC_14_0cm	1.108	2.1856	1.0776	1.9712	0.8632	19.9
	EQC_14_10cm	1.1022	2.2181	1.1159	2.1601	1.0579	5.2
	EQC_15_0cm	1.0206	2.1376	1.117	1.9083	0.8877	20.5
	EQC_15_10cm	1.1065	2.2212	1.1147	2.1374	1.0309	7.5
Mt Sugarloaf	CAS_1_0cm	1.0413	2.043	1.0017	1.9579	0.9166	8.5
	CAS_1_10cm	1.0633	2.1608	1.0975	2.1006	1.0373	5.5
	CAS_2_0cm	1.0289	2.0972	1.0683	1.9933	0.9644	9.7
	CAS_2_10cm	1.064	2.1227	1.0587	2.0063	0.9423	11.0

Supplementary Table 10 | Gas Chromatography measurements of field gas for surface (0 cm) and depth (10, 30 or 50 cm). H_2 and CO measured on the reduced compound photometer (RCP), CH_4 and NMHC on the flame ionising detector (FID), and percentages of H2, O2, N2, CH4 and CO2 measured on the thermal conductivity detector (TCD) (methods in Section 2.2.2). NA = not applicable.

	RCP			FID		TCD				
	H_2	CO	CO	CH_4	NMHC	H_2	O_2	N_2	CH_4	CO_2
	ppm	ppb	ppm	ppm	ppm	%	%	%	%	%
Air	2.9	30	0.03	0.5	6.7	0.00	21.18	72.80	0.00	0.25
Mt. Urchin	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
South Crater 0 cm	5.4	274.5	0.2745	1.3	0	0.00	21.36	73.29	0.00	0.14
South Crater 10 cm	0.9	30.2	0.0302	8.5	3	0.00	21.33	73.28	0.00	0.07
South Crater 50 cm	0.9	174.3	0.1743	0.6	0	0.00	21.42	73.49	0.00	0.05
Rec Crater 0cm	3.4	344.2	0.3442	1.7	7.1	0.00	21.39	73.50	0.00	0.05
Red Crater 10cm	3.4	169.7	0.1697	41.4	1.1	0.00	21.00	72.61	0.00	1.19
Red Crater 50cm	2.3	149.9	0.1499	186.9	0.5	0.00	20.92	72.51	0.00	1.46
Central Crater 0cm	8.4	487.7	0.4877	1.9	0.8	0.00	21.39	73.37	0.00	0.10
Central Crater 10cm	2.1	228.7	0.2287	1	5.3	0.00	21.41	73.40	0.00	0.07
Central Crater 50cm	4.6	740.7	0.7407	0	1.7	0.00	21.42	73.51	0.00	0.04
White Island 0cm	3.6	679	0.679	1.5	1.8	0.00	21.59	74.21	0.00	0.13
White Island 10cm	16.1	295.1	0.2951	8.6	10.1	0.00	19.83	68.54	0.00	6.64
White Island 50 cm	29	516.8	0.5168	18.4	/	0.00	17.87	61.99	0.00	15.07
Mt. Sugarloaf 1 0cm	5.80	275.30	0.2753	1.20	0.00	0.00	24.91	84.23	0.00	0.88
Mt. Sugarloaf 1 10cm	5.70	865.30	0.8653	0.50	4.70	0.00	25.17	85.10	0.00	0.34
Mt. Sugarloaf 1 20cm	4.50	671.30	0.6713	0.10	3.70	0.00	25.31	85.39	0.00	0.17
Mt. Sugarloaf 1 30cm	8.50	666.90	0.6669	0.80	1.70	0.00	25.20	84.95	0.00	0.13
CAS_2_0cm	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CAS_2_10cm	3.80	870.10	0.8701	1.20	0.00	0.00	25.38	85.43	0.00	0.07
CAS_2_20cm	12.10	558.40	0.5584	0.00	1.30	0.00	25.30	85.37	0.00	0.06
CAS_2_30cm	2.60	492.30	0.4923	1.10	600.00	0.00	25.33	85.47	0.00	0.06

References

1 Ji, M. *et al.* Atmospheric trace gases support primary production in Antarctic desert surface soil. *Nature*, doi:10.1038/nature25014 https://www.nature.com/articles/nature25014#supplementary-information (2017).

2 Costello, E. K., Halloy, S. R. P., Reed, S. C., Sowell, P. & Schmidt, S. K. Fumarole-Supported Islands of Biodiversity within a Hyperarid, High-Elevation Landscape on Socompa Volcano, Puna de Atacama, Andes. *Applied and Environmental Microbiology* **75**, 735-747, doi:10.1128/aem.01469-08 (2009).

3 King, G. M. Contributions of Atmospheric CO and Hydrogen Uptake to Microbial Dynamics on Recent Hawaiian Volcanic Deposits. *Applied and Environmental Microbiology* **69**, 4067-4075, doi:10.1128/aem.69.7.4067-4075.2003 (2003).

4 Lynch, R. C. *et al.* The potential for microbial life in the highest-elevation (>6000 m.a.s.l.) mineral soils of the Atacama region. *Journal of Geophysical Research: Biogeosciences* **117**, doi:doi:10.1029/2012JG001961 (2012).

5 Kósik, S., Németh, K., Lexa, J. & Procter, J. N. Understanding the evolution of a small-volume silicic fissure eruption: Puketerata Volcanic Complex, Taupo Volcanic Zone, New Zealand. *Journal of Volcanology and Geothermal Research*, doi:https://doi.org/10.1016/j.jvolgeores.2017.12.008 (2017).

6 Scott, B. J. & Potter, S. H. Aspects of historical eruptive activity and volcanic unrest at Mt. Tongariro, New Zealand: 1846–2013. *Journal of Volcanology and Geothermal Research* **286**, 263-276, doi:https://doi.org/10.1016/j.jvolgeores.2014.04.003 (2014).

7 Rothschild, L. J. & Mancinelli, R. L. Life in extreme environments. *Nature* **409**, 1092, doi:10.1038/35059215 (2001).

8 Rampelotto, P. H. Extremophiles and extreme environments. *Life (Basel, Switzerland)* **3**, 482-485, doi:10.3390/life3030482 (2013).

9 Thiel, V. in *Encyclopedia of Geobiology* (eds Joachim Reitner & Volker Thiel) 362-366 (Springer Netherlands, 2011).

10 Schulze-Makuch, D. *et al.* Transitory microbial habitat in the hyperarid Atacama Desert. **115**, 2670-2675 (2018).

11 Vincent, W. F. *et al.* Extreme ecosystems and geosystems in the Canadian High Arctic: Ward Hunt Island and vicinity. *Écoscience* **18**, 236-261, doi:10.2980/18-3-3448 (2011).

12 Baross, J. A. & Hoffman, S. E. Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life. *Origins of life and evolution of the biosphere* **15**, 327-345, doi:10.1007/BF01808177 (1985).

13 Jolivet, E. *et al.* Thermococcus gammatolerans sp. nov., a hyperthermophilic archaeon from a deep-sea hydrothermal vent that resists ionizing radiation. **53**, 847-851, doi:doi:10.1099/ijs.0.02503-0 (2003).

14 Kamekura, M. Diversity of extremely halophilic bacteria. *Extremophiles* **2**, 289-295, doi:10.1007/s007920050071 (1998).

15 Stetter, K. O. Extremophiles and their adaptation to hot environments. *FEBS Letters* **452**, 22-25, doi:doi:10.1016/S0014-5793(99)00663-8 (1999).

16 Remus-Emsermann, M. N. P. & Schlechter, R. O. Phyllosphere microbiology: at the interface between microbial individuals and the plant host. *New Phytologist* **218**, 1327-1333, doi:doi:10.1111/nph.15054 (2018).

17 Cid, F. P. *et al.* Draft genome sequences of bacteria isolated from the Deschampsia antarctica phyllosphere. *Extremophiles* **22**, 537-552, doi:10.1007/s00792-018-1015-x (2018).
18 Govil, T., Rathinam, N. K., Salem, D. R. & Sani, R. K. in *Microbial Diversity in the Genomic Era* (eds Surajit Das & Hirak Ranjan Dash) 631-656 (Academic Press, 2019).
19 Francis, P. W. in *Geochemistry* 658-659 (Springer Netherlands, 1998).
20 van Thienen, P. *et al.* Water, Life, and Planetary Geodynamical Evolution. *Space Science Reviews* **129**, 167-203, doi:10.1007/s11214-007-9149-7 (2007).
21 Simmons, S. F., Browne, P. R. L. & Brathwaite, R. L. *Active and Extinct Hydrothermal Systems of the North Island, New Zealand*. (1992).
22 Giggenbach, W. F. in *Monitoring and Mitigation of Volcano Hazards* (eds Roberto Scarpa & Robert I. Tilling) 221-256 (Springer Berlin Heidelberg, 1996).
23 Henley, R. W. & Ellis, A. J. Geothermal systems ancient and modern: a geochemical review. *Earth-Science Reviews* **19**, 1-50, doi:https://doi.org/10.1016/0012-8252(83)90075-2 (1983).
24 Stewart, L., Stucker, V., Stott, M. & Ronde, C. Marine-influenced microbial communities inhabit terrestrial hot springs on a remote island volcano. *Microbial Life Under Extreme Conditions* **22**, 687-698, doi:10.1007/s00792-018-1029-4 (2018).
25 Elder, J. W. Geothermal Systems: Principles and Case Histories, edited by L. Rybach and L. J. P. Muffler. Wiley, Chichester, 1981. No. of pages: xiv + 359 incl. index. Price £24·00; $57·00. **17**, 243-244, doi:doi:10.1002/gj.3350170305 (1982).
26 Wilson, C. J. N. & Rowland, J. V. The volcanic, magmatic and tectonic setting of the Taupo Volcanic Zone, New Zealand, reviewed from a geothermal perspective. *Geothermics* **59**, 168-187, doi:https://doi.org/10.1016/j.geothermics.2015.06.013 (2016).
27 Hobden, B. J., Houghton, B. F., Davidson, J. P. & Weaver, S. D. Small and short-lived magma batches at composite volcanoes: time windows at Tongariro volcano, New Zealand. **156**, 865-868, doi:10.1144/gsjgs.156.5.0865 %J Journal of the Geological Society (1999).
28 Wilson, C., Gravley, D., Leonard, G. & Rowland, J. J. S. i. v. t. l. o. G. W. S. P. o. I. Volcanism in the central Taupo Volcanic Zone, New Zealand: tempo, styles and controls. **2**, 225-247 (2009).
29 Wilson, C. J. N. *et al.* Volcanic and structural evolution of Taupo Volcanic Zone, New Zealand: a review. *Journal of Volcanology and Geothermal Research* **68**, 1-28, doi:https://doi.org/10.1016/0377-0273(95)00006-G (1995).
30 Cole, J. W. J. B. o. V. Structural control and origin of volcanism in the Taupo volcanic zone, New Zealand. **52**, 445-459, doi:10.1007/bf00268925 (1990).
31 Hobden, B. J., Houghton, B. F. & Nairn, I. A. Growth of a young, frequently active composite cone: Ngauruhoe volcano, New Zealand. *Bulletin of Volcanology* **64**, 392-409, doi:10.1007/s00445-002-0216-3 (2002).
32 Walsh, F., Hochstein, M. & Bromley, C. in *Proceedings 20th Geothermal Workshop.* 317-324.
33 Scott, B. J. & Travers, J. J. N. h. Volcano monitoring in NZ and links to SW Pacific via the Wellington VAAC. **51**, 263-273 (2009).
34 Jolly, A. D., Sherburn, S., Jousset, P. & Kilgour, G. Eruption source processes derived from seismic and acoustic observations of the 25 September 2007 Ruapehu eruption—North Island, New Zealand. *Journal of Volcanology and Geothermal Research* **191**, 33-45, doi:https://doi.org/10.1016/j.jvolgeores.2010.01.009 (2010).
35 Bay, S., Ferrari, B. & Greening, C. *Life without water: how do bacteria generate biomass in desert ecosystems?* , (2018).

36 Lee, K. C. *et al.* Stochastic and Deterministic Effects of a Moisture Gradient on Soil Microbial Communities in the McMurdo Dry Valleys of Antarctica. *Frontiers in microbiology* **9**, 2619-2619, doi:10.3389/fmicb.2018.02619 (2018).

37 Lynch, R. C., Darcy, J. L., Kane, N. C., Nemergut, D. R. & Schmidt, S. K. Metagenomic evidence for metabolism of trace atmospheric gases by high-elevation desert Actinobacteria. *Front. Microbiol.* **5**, 698 (2014).

38 Fell, J. W., Scorzetti, G., Connell, L. & Craig, S. Biodiversity of micro-eukaryotes in Antarctic Dry Valley soils with <5% soil moisture. *Soil Biology and Biochemistry* **38**, 3107-3119, doi:10.1016/j.soilbio.2006.01.014 (2006).

39 Cary, S. C., McDonald, I. R., Barrett, J. E. & Cowan, D. A. On the rocks: the microbiology of Antarctic Dry Valley soils. *Nature Reviews Microbiology* **8**, 129, doi:10.1038/nrmicro2281 (2010).

40 Cockell, C. S., Kelly, L. C. & Marteinsson, V. Actinobacteria—An Ancient Phylum Active in Volcanic Rock Weathering. *Geomicrobiology Journal* **30**, 706-720, doi:10.1080/01490451.2012.758196 (2013).

41 Abdulla, H. Bioweathering and biotransformation of granitic rock minerals by actinomycetes. *Microb Ecol* **58**, 753-761, doi:10.1007/s00248-009-9549-1 (2009).

42 Battistuzzi, F. U. & Hedges, S. B. A major clade of prokaryotes with ancient adaptations to life on land. *Molecular biology and evolution* **26**, 335-343, doi:10.1093/molbev/msn247 (2009).

43 King, G. M. Molecular and Culture-Based Analyses of Aerobic Carbon Monoxide Oxidizer Diversity†. *Applied and Environmental Microbiology* **69**, 7257-7265, doi:10.1128/AEM.69.12.7257-7265.2003 (2003).

44 Greening, C. *et al.* Atmospheric Hydrogen Scavenging: from Enzymes to Ecosystems. *Applied and Environmental Microbiology* **81**, 1190-1199, doi:10.1128/aem.03364-14 (2015).

45 Islam, Z. F. *et al.* Two Chloroflexi classes independently evolved the ability to persist on atmospheric hydrogen and carbon monoxide. *bioRxiv*, 457697, doi:10.1101/457697 (2018).

46 King, G. M. & Weber, C. F. Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nat. Rev. Microbiol.* **5**, 107-118 (2007).

47 Janssen, P. H. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl Environ Microbiol* **72**, 1719-1728, doi:10.1128/aem.72.3.1719-1728.2006 (2006).

48 Bergmann, G. T. *et al.* The under-recognized dominance of Verrucomicrobia in soil bacterial communities. *Soil Biol Biochem* **43**, 1450-1455, doi:10.1016/j.soilbio.2011.03.012 (2011).

49 King, A. J. *et al.* Biogeography and habitat modelling of high-alpine bacteria. *Nature Communications* **1**, 53, doi:10.1038/ncomms1055 https://www.nature.com/articles/ncomms1055#supplementary-information (2010).

50 Lee, K. C. Y. *et al.* Phylogenetic Delineation of the Novel Phylum Armatimonadetes (Former Candidate Division OP10) and Definition of Two Novel Candidate Divisions. **79**, 2484-2487, doi:10.1128/AEM.03333-12 %J Applied and Environmental Microbiology (2013).

51 Rinke, C. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**, 431-437 (2013).

52 Power, J. F. *et al.* Microbial biogeography of 925 geothermal springs in New Zealand. *Nature Communications* **9**, 2876, doi:10.1038/s41467-018-05020-y (2018).

53 Oliverio, A. M. *et al.* The ecology and diversity of microbial eukaryotes in geothermal springs. *The ISME journal* **12**, 1918-1928, doi:10.1038/s41396-018-0104-2 (2018).
54 Stetter, K. O. Hyperthermophiles : Isolation, classification, and properties, in Extremophiles. *Microbial Life in Extreme Environments*, 1-24 (1998).
55 Horikoshi, K. Barophiles: deep-sea microorganisms adapted to an extreme environment. *Current opinion in microbiology* **1**, 291-295 (1998).
56 Martin, W., Baross, J., Kelley, D. & Russell, M. J. J. N. R. M. Hydrothermal vents and the origin of life. **6**, 805 (2008).
57 Stetter, K. O. Hyperthermophiles in the history of life. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **361**, 1837-1843, doi:10.1098/rstb.2006.1907 (2006).
58 Satyanarayana, T., Littlechild, J. & Kawarabayasi, Y. in *biotechnology of thermophiles.* (Springer, 2013).
59 Amarouche-Yala, S., Benouadah, A., El Ouahab Bentabet, A. & Lopez-Garcia, P. Morphological and phylogenetic diversity of thermophilic cyanobacteria in Algerian hot springs. *Extremophiles* **18**, 1035-1047, doi:10.1007/s00792-014-0680-7 (2014).
60 Vetriani, C., Speck, M. D., Ellor, S. V., Lutz, R. A. & Starovoytov, V. Thermovibrio ammonificans sp. nov., a thermophilic, chemolithotrophic, nitrate-ammonifying bacterium from deep-sea hydrothermal vents. *Int J Syst Evol Microbiol* **54**, 175-181, doi:10.1099/ijs.0.02781-0 (2004).
61 Quatrini, R. & Johnson, D. B. Microbiomes in extremely acidic environments: functionalities and interactions that allow survival and growth of prokaryotes at low pH. *Current opinion in microbiology* **43**, 139-147, doi:https://doi.org/10.1016/j.mib.2018.01.011 (2018).
62 Khadem, A. F. *et al.* Genomic and Physiological Analysis of Carbon Storage in the Verrucomicrobial Methanotroph "Ca. Methylacidiphilum Fumariolicum" SolV. *Frontiers in microbiology* **3**, 345-345, doi:10.3389/fmicb.2012.00345 (2012).
63 Carere, C. R. *et al.* Mixotrophy drives niche expansion of verrucomicrobial methanotrophs. *The ISME journal* **11**, 2599, doi:10.1038/ismej.2017.112 https://www.nature.com/articles/ismej2017112#supplementary-information (2017).
64 Grimbergen, A. J., Siebring, J., Solopova, A. & Kuipers, O. P. Microbial bet-hedging: the power of being different. *Current opinion in microbiology* **25**, 67-72, doi:https://doi.org/10.1016/j.mib.2015.04.008 (2015).
65 Berney, M., Greening, C., Hards, K., Collins, D. & Cook, G. M. Three different [NiFe] hydrogenases confer metabolic flexibility in the obligate aerobe Mycobacterium smegmatis. *Environ. Microbiol.* **16**, 318-330 (2014).
66 Greening, C. *et al.* Genomic and metagenomic surveys of hydrogenase distribution indicate H2 is a widely utilised energy source for microbial growth and survival. *The ISME journal* **10**, 761-777, doi:10.1038/ismej.2015.153 (2016).
67 Logan, J. A., Prather, M. J., Wofsy, S. C. & McElroy, M. B. Tropospheric chemistry: A global perspective. *Journal of Geophysical Research: Oceans* **86**, 7210-7254, doi:doi:10.1029/JC086iC08p07210 (1981).
68 Monson, R. K. & Holland, E. A. Biospheric Trace Gas Fluxes and Their Control Over Tropospheric Chemistry. *Annual Review of Ecology and Systematics* **32**, 547-576, doi:10.1146/annurev.ecolsys.32.081501.114136 (2001).
69 Stubbins, A. *et al.* Open-ocean carbon monoxide photoproduction. *Deep Sea Research Part II: Topical Studies in Oceanography* **53**, 1695-1705, doi:https://doi.org/10.1016/j.dsr2.2006.05.011 (2006).

70 Martínez-Alonso, S. *et al.* First satellite identification of volcanic carbon monoxide. *Geophysical Research Letters* **39**, doi:doi:10.1029/2012GL053275 (2012).
71 Schmidt, U. & Conrad, R. Hydrogen, carbon monoxide, and methane dynamics in Lake Constance. *Limnology and Oceanography* **38**, 1214-1226, doi:doi:10.4319/lo.1993.38.6.1214 (1993).
72 Manning, M. R., Lowe, D. C., Moss, R. C., Bodeker, G. E. & Allan, W. Short-term variations in the oxidizing power of the atmosphere. *Nature* **436**, 1001-1004, doi:10.1038/nature03900 (2005).
73 Cody, G. D. *et al.* Primordial carbonylated iron-sulfur compounds and the synthesis of pyruvate. *Science* **289**, 1337-1340 (2000).
74 Aylward, N. & Bofinger, N. The reactions of methanimine and cyanogen with carbon monoxide in prebiotic molecular evolution on earth. *Origins of life and evolution of the biosphere : the journal of the International Society for the Study of the Origin of Life* **31**, 481-500 (2001).
75 Miyakawa, S., Yamanashi, H., Kobayashi, K., Cleaves, H. J. & Miller, S. L. Prebiotic synthesis from CO atmospheres: Implications for the origins of life. *Proceedings of the National Academy of Sciences* **99**, 14628-14631, doi:10.1073/pnas.192568299 (2002).
76 Crutzen, P. J. & Gidel, L. T. A two-dimensional photochemical model of the atmosphere: 2. The tropospheric budgets of the anthropogenic chlorocarbons CO, CH4, CH3Cl and the effect of various NO x sources on tropospheric ozone. *Journal of Geophysical Research: Oceans* **88**, 6641-6661, doi:doi:10.1029/JC088iC11p06641 (1983).
77 Holloway, T., Levy, H. & Kasibhatla, P. Global distribution of carbon monoxide. *Journal of Geophysical Research: Atmospheres* **105**, 12123-12147, doi:doi:10.1029/1999JD901173 (2000).
78 Pétron, G. *et al.* Monthly CO surface sources inventory based on the 2000–2001 MOPITT satellite data. *Geophysical Research Letters* **31**, doi:doi:10.1029/2004GL020560 (2004).
79 Lelieveld, J., Dentener, F. J., Peters, W. & Krol, M. C. On the role of hydroxyl radicals in the self-cleansing capacity of the troposphere. *Atmos. Chem. Phys.* **4**, 2337-2344, doi:10.5194/acp-4-2337-2004 (2004).
80 Dias, J. M. *et al.* Crystal structure of the first dissimilatory nitrate reductase at 1.9 A solved by MAD methods. *Structure (London, England : 1993)* **7**, 65-79 (1999).
81 Lu, Y. & Khalil, M. A. K. Methane and carbon monoxide in OH chemistry: The effects of feedbacks and reservoirs generated by the reactive products. *Chemosphere* **26**, 641-655, doi:https://doi.org/10.1016/0045-6535(93)90450-J (1993).
82 Daniel, J. S. & Solomon, S. On the climate forcing of carbon monoxide. *Journal of Geophysical Research: Atmospheres* **103**, 13249-13260, doi:doi:10.1029/98JD00822 (1998).
83 Crutzen, P. J. The Role of NO and NO2 in the Chemistry of the Troposphere and Stratosphere. *Annual Review of Earth and Planetary Sciences* **7**, 443-472, doi:10.1146/annurev.ea.07.050179.002303 (1979).
84 Greening, C., Berney, M., Hards, K., Cook, G. M. & Conrad, R. A soil actinobacterium scavenges atmospheric H(2) using two membrane-associated, oxygen-dependent [NiFe] hydrogenases. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 4257-4261, doi:10.1073/pnas.1320586111 (2014).
85 Constant, P., Poissant, L. & Villemur, R. Tropospheric H2 budget and the response of its soil uptake under the changing environment. *Science of The Total*

Environment **407**, 1809-1823, doi:https://doi.org/10.1016/j.scitotenv.2008.10.064 (2009).

86 Seinfeld, J., Pandis, S. J. A. C. & Change, P. F. A. P. t. C. Chemistry of the atmospheric aqueous phase. **337407** (1998).

87 Schwartz, E., Fritsch, J. & Friedrich, B. in *The Prokaryotes: Prokaryotic Physiology and Biochemistry* (eds Eugene Rosenberg *et al.*) 119-199 (Springer Berlin Heidelberg, 2013).

88 Sokolova, T. G. *et al.* The first evidence of anaerobic CO oxidation coupled with H2 production by a hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. *Extremophiles* **8**, 317-323, doi:10.1007/s00792-004-0389-0 (2004).

89 Dixon, R. O. D. J. A. f. M. Hydrogenase in legume root nodule bacteroids: Occurrence and properties. **85**, 193-201, doi:10.1007/bf00408844 (1972).

90 Schmidt, U. Molecular hydrogen in the atmosphere. **26**, 78-90, doi:doi:10.1111/j.2153-3490.1974.tb01954.x (1974).

91 Dick, G. J. *et al.* The microbiology of deep-sea hydrothermal vent plumes: ecological and biogeographic linkages to seafloor and water column habitats. *Frontiers in microbiology* **4**, 124-124, doi:10.3389/fmicb.2013.00124 (2013).

92 Kolb, S., Horn, M. A., Murrell, J. C. & Knief, C. Editorial: The Impact of Microorganisms on Consumption of Atmospheric Trace Gases. *Frontiers in Microbiology* **8**, doi:10.3389/fmicb.2017.01856 (2017).

93 Conrad, R. Soil microorganisms oxidizing atmospheric trace gases (CH4, CO, H2, NO). *Indian J. Microbiol.* **39**, 193-203 (1999).

94 Sokolova, T. G. *et al.* Aerobic carbon monoxide oxidation in the course of growth of a hyperthermophilic archaeon, Sulfolobus sp. ETSY. *Microbiology (Reading, England)* **86**, 539-548, doi:10.1134/s0026261717050174 (2017).

95 Svetlichny, V. A. *et al.* Carboxydothermus hydrogenoformans gen. nov., sp. nov., a CO-utilizing Thermophilic Anaerobic Bacterium from Hydrothermal Environments of Kunashir Island. *Systematic and Applied Microbiology* **14**, 254-260, doi:https://doi.org/10.1016/S0723-2020(11)80377-2 (1991).

96 Klenk, H. P. *et al.* The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon Archaeoglobus fulgidus. *Nature* **390**, 364-370, doi:10.1038/37052 (1997).

97 Wu, M. *et al.* Life in hot carbon monoxide: the complete genome sequence of Carboxydothermus hydrogenoformans Z-2901. *PLoS genetics* **1**, e65, doi:10.1371/journal.pgen.0010065 (2005).

98 Dobbek, H., Gremer, L., Meyer, O. & Huber, R. Crystal structure and mechanism of CO dehydrogenase, a molybdo iron-sulfur flavoprotein containing S-selanylcysteine. *Proceedings of the National Academy of Sciences* **96**, 8884-8889, doi:10.1073/pnas.96.16.8884 (1999).

99 Conrad, R. Soil microorganisms as controllers of atmospheric trace gases (H2, CO, CH4, OCS, N2O, and NO. *Microbiological reviews* **60** (1996).

100 Rohde, M., Mayer, F. & Meyer, O. Immunocytochemical localization of carbon monoxide oxidase in Pseudomonas carboxydovorans. The enzyme is attached to the inner aspect of the cytoplasmic membrane. *Journal of Biological Chemistry* **259**, 14788-14792 (1984).

101 Hausinger, R. P. Metallocenter assembly in nickel-containing enzymes. *JBIC Journal of Biological Inorganic Chemistry* **2**, 279-286, doi:10.1007/s007750050133 (1997).

102 Adam, P. S., Borrel, G. & Gribaldo, S. Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. *Proceedings of the National Academy of Sciences* **115**, E1166-E1173, doi:10.1073/pnas.1716667115 (2018).

103 Forster, P. *et al. Changes in Atmospheric Constituents and in Radiative Forcing Chapter 2*. (Cambridge University Press, 2007).
104 Santiago, B., Schubel, U., Egelseer, C. & Meyer, O. Sequence analysis, characterization and CO-specific transcription of the cox gene cluster on the megaplasmid pHCG3 of Oligotropha carboxidovorans. *Gene* **236**, 115-124 (1999).
105 Hille, R. Molybdenum-containing hydroxylases. *Archives of biochemistry and biophysics* **433**, 107-116, doi:10.1016/j.abb.2004.08.012 (2005).
106 Fuhrmann, S. *et al.* Complete nucleotide sequence of the circular megaplasmid pHCG3 of Oligotropha carboxidovorans: function in the chemolithoautotrophic utilization of CO, H(2) and CO(2). *Gene* **322**, 67-75 (2003).
107 Wu, D. *et al.* Complete genome sequence of the aerobic CO-oxidizing thermophile Thermomicrobium roseum. *PloS one* **4**, e4207, doi:citeulike-article-id:3893307

doi: 10.1371/journal.pone.0004207 (2009).
108 Lorite, M. J., Tachil, J., Sanjuan, J., Meyer, O. & Bedmar, E. J. Carbon monoxide dehydrogenase activity in Bradyrhizobium japonicum. *Appl Environ Microbiol* **66**, 1871-1876 (2000).
109 Ehhalt, D. H. & Rohrer, F. The tropospheric cycle of H2: a critical review. **61**, 500-535, doi:doi:10.1111/j.1600-0889.2009.00416.x (2009).
110 Ehhalt, D. Gas phase chemistry of the troposphere. *Global aspects of atmospheric chemistry* **6**, 21-110 (1999).
111 Kaserer, H. *Die Oxydation des Wasserstoffs durch Mikroorganismen*, 1906).
112 Schlegel, H. G. The physiology of hydrogen bacteria. *Antonie van Leeuwenhoek* **42**, 181-201, doi:10.1007/BF00394115 (1976).
113 Constant, P., Poissant, L. & Villemur, R. Isolation of Streptomyces sp. PCB7, the first microorganism demonstrating high-affinity uptake of tropospheric H2. *The ISME journal* **2**, 1066-1076, doi:10.1038/ismej.2008.59 (2008).
114 Constant, P., Chowdhury, S. P., Pratscher, J. & Conrad, R. Streptomycetes contributing to atmospheric molecular hydrogen soil uptake are widespread and encode a putative high-affinity [NiFe]-hydrogenase. *Environ. Microbiol.* **12**, 821-829 (2010).
115 Kessler, E. Effect of anaerobiosis on photosynthetic reactions and nitrogen metabolism of algae with and without hydrogenase. *Archiv fur Mikrobiologie* **93**, 91-100 (1973).
116 Lindmark, D. G. & Muller, M. Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate Tritrichomonas foetus, and its role in pyruvate metabolism. *The Journal of biological chemistry* **248**, 7724-7728 (1973).
117 Danovaro, R. *et al.* The first metazoa living in permanently anoxic conditions. *BMC biology* **8**, 30, doi:10.1186/1741-7007-8-30 (2010).
118 Yagi, T. & Higuchi, Y. Studies on hydrogenase. *Proceedings of the Japan Academy. Series B, Physical and biological sciences* **89**, 16-33, doi:10.2183/pjab.89.16 (2013).
119 Marques, M. C. *et al.* The direct role of selenocysteine in [NiFeSe] hydrogenase maturation and catalysis. *Nature Chemical Biology* **13**, 544, doi:10.1038/nchembio.2335 https://www.nature.com/articles/nchembio.2335#supplementary-information (2017).
120 Buckel, W. & Thauer, R. K. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na+ translocating ferredoxin oxidation. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1827**, 94-113, doi:https://doi.org/10.1016/j.bbabio.2012.07.002 (2013).
121 Ogata, H., Lubitz, W. & Higuchi, Y. Structure and function of [NiFe] hydrogenases. *The Journal of Biochemistry* **160**, 251-258, doi:10.1093/jb/mvw048 (2016).

122 Liebgott, P.-P. *et al.* Relating diffusion along the substrate tunnel and oxygen sensitivity in hydrogenase. *Nature Chemical Biology* **6**, 63, doi:10.1038/nchembio.276 https://www.nature.com/articles/nchembio.276#supplementary-information (2009).

123 Burgdorf, T. *et al.* [NiFe]-hydrogenases of Ralstonia eutropha H16: modular enzymes for oxygen-tolerant biological hydrogen oxidation. *Journal of molecular microbiology and biotechnology* **10**, 181-196, doi:10.1159/000091564 (2005).

124 Greening, C. *et al.* Alternative hydrogen uptake pathways suppress methane production in ruminants. *bioRxiv*, 486894, doi:10.1101/486894 (2018).

125 Mulder, D. W. *et al.* Stepwise [FeFe]-hydrogenase H-cluster assembly revealed in the structure of HydAΔEFG. *Nature* **465**, 248, doi:10.1038/nature08993 https://www.nature.com/articles/nature08993#supplementary-information (2010).

126 Berggren, G. *et al.* Biomimetic assembly and activation of [FeFe]-hydrogenases. *Nature* **499**, 66-69, doi:10.1038/nature12239 (2013).

127 Martiny, J. B. H. *et al.* Microbial biogeography: putting microorganisms on the map. *Nat Rev Micro* **4**, 102-112 (2006).

128 van der Gast, C. J. Microbial biogeography: the end of the ubiquitous dispersal hypothesis? *Environmental Microbiology* **17**, 544-546, doi:10.1111/1462-2920.12635 (2015).

129 Zheng, Y.-M., Cao, P., Fu, B., Hughes, J. M. & He, J.-Z. Ecological Drivers of Biogeographic Patterns of Soil Archaeal Community. *PLoS ONE* **8**, e63375, doi:10.1371/journal.pone.0063375 (2013).

130 Xu, J. Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. *Mol Ecol* **15**, 1713-1731, doi:10.1111/j.1365-294X.2006.02882.x (2006).

131 Stefani, F. O. P. *et al.* Culture-Dependent and -Independent Methods Capture Different Microbial Community Fractions in Hydrocarbon-Contaminated Soils. *PLoS ONE* **10**, e0128272, doi:10.1371/journal.pone.0128272 (2015).

132 Turnbaugh, P. J. *et al.* The human microbiome project. **449**, 804 (2007).

133 Park, S. T. & Kim, J. Trends in Next-Generation Sequencing and a New Era for Whole Genome Sequencing. *International Neurourology Journal* **20**, S76-83, doi:10.5213/inj.1632742.371 (2016).

134 Wheeler, D. L. *et al.* Database resources of the National Center for Biotechnology Information. *Nucleic acids research* **36**, D13-D21, doi:10.1093/nar/gkm1000 (2008).

135 Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* **10**, 996-998, doi:10.1038/nmeth.2604 (2013).

136 Gutleben, J. *et al.* The multi-omics promise in context: from sequence to microbial isolate. *Critical reviews in microbiology* **44**, 212-229, doi:10.1080/1040841x.2017.1332003 (2018).

137 Chu, H. *et al.* Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. *Environmental Microbiology* **12**, 2998-3006, doi:10.1111/j.1462-2920.2010.02277.x (2010).

138 Niederberger, T. D. *et al.* Microbial community composition of transiently wetted Antarctic Dry Valley soils. *Frontiers in Microbiology* **6**, doi:10.3389/fmicb.2015.00009 (2015).

139 Gilbert, J. A. *et al.* Current understanding of the human microbiome. *Nature Medicine* **24**, 392, doi:10.1038/nm.4517 (2018).

140 Aguinaga, O. E., McMahon, A., White, K. N., Dean, A. P. & Pittman, J. K. Microbial Community Shifts in Response to Acid Mine Drainage Pollution Within a Natural

Wetland Ecosystem. *Frontiers in Microbiology* **9**, doi:10.3389/fmicb.2018.01445 (2018).

141 Griffiths, R. I. *et al.* The bacterial biogeography of British soils. *Environ Microbiol* **13**, 1642-1654, doi:10.1111/j.1462-2920.2011.02480.x (2011).

142 Bahl, J. *et al.* Ancient origins determine global biogeography of hot and cold desert cyanobacteria. *Nature Communications* **2**, 163, doi:10.1038/ncomms1167 http://www.nature.com/articles/ncomms1167#supplementary-information (2011).

143 Pointing, S. B. *et al.* Highly specialized microbial diversity in hyper-arid polar desert. *Proceedings of the National Academy of Sciences* **106**, 19964-19969, doi:10.1073/pnas.0908274106 (2009).

144 Anderson, R. E., Sogin, M. L. & Baross, J. A. Biogeography and ecology of the rare and abundant microbial lineages in deep-sea hydrothermal vents. *FEMS Microbiol Ecol* **91**, 1-11, doi:10.1093/femsec/fiu016 (2015).

145 Wu, J., Anderson, B. J., Buckley, H. L., Lewis, G. & Lear, G. Aspect has a greater impact on alpine soil bacterial community structure than elevation. *FEMS Microbiology Ecology* **93**, fiw253-fiw253, doi:10.1093/femsec/fiw253 (2017).

146 Fujimura, R. *et al.* Analysis of Early Bacterial Communities on Volcanic Deposits on the Island of Miyake (Miyake-jima), Japan: a 6-year Study at a Fixed Site. *Microbes and Environments* **27**, 19-29, doi:10.1264/jsme2.ME11207 (2012).

147 Weber, C. F. & King, G. M. Volcanic Soils as Sources of Novel CO-Oxidizing Paraburkholderia and Burkholderia: Paraburkholderia hiiakae sp. nov., Paraburkholderia metrosideri sp. nov., Paraburkholderia paradisi sp. nov., Paraburkholderia peleae sp. nov., and Burkholderia alpina sp. nov. a Member of the Burkholderia cepacia Complex. *Frontiers in Microbiology* **8**, doi:10.3389/fmicb.2017.00207 (2017).

148 Lee, K. C. *et al.* Niche Filtering of Bacteria in Soil and Rock Habitats of the Colorado Plateau Desert, Utah, USA. *Frontiers in Microbiology* **7**, doi:10.3389/fmicb.2016.01489 (2016).

149 Kent, A. G., Dupont, C. L., Yooseph, S. & Martiny, A. C. Global biogeography of Prochlorococcus genome diversity in the surface ocean. *The ISME journal* **10**, 1856-1865, doi:10.1038/ismej.2015.265 (2016).

150 Salazar, G. *et al.* Global diversity and biogeography of deep-sea pelagic prokaryotes. *The ISME journal* **10**, 596-608, doi:10.1038/ismej.2015.137 (2016).

151 Gilbert, J. A. & Stephens, B. Microbiology of the built environment. *Nature Reviews Microbiology* **16**, 661-670, doi:10.1038/s41579-018-0065-5 (2018).

152 Goldford, J. E. *et al.* Emergent simplicity in microbial community assembly. *Science* **361**, 469 (2018).

153 Nemergut, D. R. *et al.* Patterns and Processes of Microbial Community Assembly. *Microbiology and molecular biology reviews : MMBR* **77**, 342-356, doi:10.1128/MMBR.00051-12 (2013).

154 Livermore, J. A. & Jones, S. E. Local–global overlap in diversity informs mechanisms of bacterial biogeography. *The ISME journal* **9**, 2413-2422, doi:10.1038/ismej.2015.51 (2015).

155 Philipson, W. *et al.* Transactions and Proceedings of the Royal Society of New Zealand 1868-1961. (1961).

156 Thorn, V. C. Vegetation reconstruction from soil phytoliths, Tongariro National Park, New Zealand. *New Zealand Journal of Botany* **44**, 397-413, doi:10.1080/0028825X.2006.9513031 (2006).

157 Stott, M. B. *et al.* Isolation of novel bacteria, including a candidate division, from geothermal soils in New Zealand. *Environ Microbiol* **10**, 2030-2041, doi:10.1111/j.1462-2920.2008.01621.x (2008).

158 Rhodes, M. *et al.* in *The Prokaryotes* 43-55 (Springer, 2013).
159 Fierer, N. *et al.* Reconstructing the Microbial Diversity and Function of Pre-Agricultural Tallgrass Prairie Soils in the United States. *Science* **342**, 621, doi:10.1126/science.1243768 (2013).
160 Reynolds, S. The gravimetric method of soil moisture determination. *Journal of Hydrology* **11**, 258-273 (1970).
161 Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**, 335, doi:10.1038/nmeth.f.303 https://www.nature.com/articles/nmeth.f.303#supplementary-information (2010).
162 Schloss, P. D. *et al.* Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**, doi:10.1128/aem.01541-09 (2009).
163 Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* **17**, 3 %J EMBnet.journal, doi:10.14806/ej.17.1.200 (2011).
164 Wang, Q., Garrity, G. M., Tiedje, J. M., Cole, J. R. J. A. & microbiology, e. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. **73**, 5261-5267 (2007).
165 Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41**, D590-D596, doi:10.1093/nar/gks1219 (2013).
166 McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**, e61217, doi:10.1371/journal.pone.0061217 (2013).
167 Oksanen, A. *et al.* (2018).
168 Wickham, H. in *ggplot2* 3-10 (Springer, 2016).
169 Aho, K., Derryberry, D. & Peterson, T. Model selection for ecologists: the worldviews of AIC and BIC. **95**, 631-636, doi:doi:10.1890/13-1452.1 (2014).
170 Dormann, C. F. *et al.* Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. **36**, 27-46, doi:doi:10.1111/j.1600-0587.2012.07348.x (2013).
171 Caporaso, J. G. *et al.* PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **26**, 266-267, doi:10.1093/bioinformatics/btp636 (2010).
172 Price, M., Dehal, P. & Arkin, A. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* **5**, e9490, doi:citeulike-article-id:6796596

doi: 10.1371/journal.pone.0009490 (2010).
173 Gardener, M. *Statistics for Ecologists Using R and Excel: Data Collection, Exploration, Analysis and Presentation*. (Pelagic Publishing, 2017).
174 Cole, J. W. Andesites of the Tongariro volcanic centre, North Island, New Zealand. *Journal of Volcanology and Geothermal Research* **3**, 121-153, doi:https://doi.org/10.1016/0377-0273(78)90007-0 (1978).
175 Ramirez, K. S. *et al.* Biogeographic patterns in below-ground diversity in New York City's Central Park are similar to those observed globally. *Proceedings. Biological sciences* **281**, doi:10.1098/rspb.2014.1988 (2014).
176 Fierer, N. *et al.* Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences* **109**, 21390-21395, doi:10.1073/pnas.1215210110 (2012).
177 Robinson, C. K. *et al.* Microbial diversity and the presence of algae in halite endolithic communities are correlated to atmospheric moisture in the hyper-arid

zone of the Atacama Desert. *Environ Microbiol* **17**, 299-315, doi:10.1111/1462-2920.12364 (2015).

178 Vignais, P. M. & Billoud, B. Occurrence, Classification, and Biological Function of Hydrogenases: An Overview. *Chemical Reviews* **107**, 4206-4272, doi:10.1021/cr050196r (2007).

179 Constant, P., Chowdhury, S. P., Hesse, L., Pratscher, J. & Conrad, R. Genome Data Mining and Soil Survey for the Novel Group 5 [NiFe]-Hydrogenase to Explore the Diversity and Ecological Importance of Presumptive High Affinity H2-Oxidizing Bacteria. *Applied and Environmental Microbiology*, doi:10.1128/aem.00673-11 (2011).

180 Upstill-Goddard, R. C. & Barnes, J. Methane emissions from UK estuaries: Re-evaluating the estuarine source of tropospheric methane from Europe. *Marine Chemistry* **180**, 14-23, doi:https://doi.org/10.1016/j.marchem.2016.01.010 (2016).

181 Quiza, L., Lalonde, I., Guertin, C. & Constant, P. Land-use influences the distribution and activity of high affinity CO-oxidizing bacteria associated to type I-coxL genotype in soil. *Front. Microbiol.* **5**, 271 (2014).

182 Sanger, F. & Coulson, A. R. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology* **94**, 441-448, doi:https://doi.org/10.1016/0022-2836(75)90213-2 (1975).

183 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* **215**, 403-410, doi:10.1016/s0022-2836(05)80360-2 (1990).

184 Kelly, L. C., Cockell, C. S., Thorsteinsson, T., Marteinsson, V. & Stevenson, J. Pioneer Microbial Communities of the Fimmvörðuháls Lava Flow, Eyjafjallajökull, Iceland. *Microbial Ecology* **68**, 504-518, doi:10.1007/s00248-014-0432-3 (2014).

185 Weber, C. F. & King, G. M. Distribution and diversity of carbon monoxide-oxidizing bacteria and bulk bacterial communities across a succession gradient on a Hawaiian volcanic deposit. *Environmental Microbiology* **12**, 1855-1867, doi:doi:10.1111/j.1462-2920.2010.02190.x (2010).

186 King, C. E. & King, G. M. Temperature responses of carbon monoxide and hydrogen uptake by vegetated and unvegetated volcanic cinders. *The ISME journal* **6**, 1558, doi:10.1038/ismej.2011.206 https://www.nature.com/articles/ismej2011206#supplementary-information (2012).

187 Picek, T., Šimek, M. & Šantrůčková, H. Microbial responses to fluctuation of soil aeration status and redox conditions. *Biology and Fertility of Soils* **31**, 315-322, doi:10.1007/s003740050662 (2000).

188 Constant, P., Chowdhury, S. P. & Conrad, R. Co-localization of atmospheric H2 oxidation activity and high affinity H2-oxidizing bacteria in non-axenic soil and sterile soil amended with Streptomyces sp. PCB7. *Biol. Biochem.* **43**, 1888-1893 (2011).

189 Greening, C. *et al.* Persistence of the dominant soil phylum Acidobacteria by trace gas scavenging. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 10497-10502, doi:10.1073/pnas.1508385112 (2015).

190 King, C. E. & King, G. M. Description of Thermogemmatispora carboxidivorans sp. nov., a carbon-monoxide-oxidizing member of the class Ktedonobacteria isolated from a geothermally heated biofilm, and analysis of carbon monoxide oxidation by members of the class Ktedonobacteria. *Int J Syst Evol Microbiol* **64**, 1244-1251, doi:10.1099/ijs.0.059675-0 (2014).

191 Carini, P. *et al.* Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. **2**, 16242, doi:10.1038/nmicrobiol.2016.242

https://www.nature.com/articles/nmicrobiol2016242#supplementary-information (2016).