Actinobacteria associated with two diverse soil environments and their multicopper oxidase diversity

Alaric Prins

3107296

UNIVERSITY *of the*
WESTERN CAPE

A thesis submitted in partial fulfilment of the requirements for the degree *Doctor Philosophae* in the Department of Biotechnology, University of the Western Cape

Supervisors

Associate Professor Marilize Le Roes-Hill
Cape Peninsula University of Technology

Associate Professor Bronwyn Kirby-McCullough
University of the Western Cape

March 2024

ABSTRACT

Actinobacteria associated with two diverse soil environments and their multicopper oxidase diversity

Alaric Prins

PhD Thesis, University of the Western Cape

The Cape Floristic Region (CFR) is a biodiverse region boasting unique plant diversity with a rich concentration of endemic plants. *Aspalathus linearis* (Rooibos) is an indigenous plant that grows in the Clanwilliam region of the Western Cape and is cultivated for its use as an herbal tea. Emerging peatlands in the CFR have gained increasing attention over recent years through research aiming to understand the microbial diversity associated with these environments. Little is known about the actinobacterial diversity of these regions, and as such, it is necessary to investigate the diversity of the actinobacteria associated with these environments, whilst simultaneously gaining knowledge on whether the associated actinobacteria may produce enzymes of biotechnological interest.

Two CFR regions (the Rooibos environment – Clanwillian, and the Springfield emerging peatland environment – Agulhas) were explored through culture-based and genomic screening. Metabarcoding analyses using actinobacterial-specific 16S rRNA gene primers showed that the major taxa contributing to the Rooibos environment were members of the families *Mycobacteriaceae, Pseudonocardiaceae, Frankiaceae* and *Geodermatophilaceae.* Members of the families *Mycobacteriaceaea, Pseudonocardiaceae, Acidimicrobiaceae* and *Nocardioiaceae* was identified as the major taxa for the Springfield environment.

Through selective isolation techniques, actinobacteria from rare (underrepresented) genera were isolated, including members of the genera *Dactylosporangium, Actinokineospora, Curtobacterium, Modestobacter, Leifsonia* and *Actinomadura.* The top strains, selected based on exhibiting extracellular multicopper oxidase (MCO) activity through culture-based screening, were subjected to whole genome sequence analysis. These rare genera are also vastly underrepresented among 3 400 bacterial MCO

sequences found in the Laccase and Multicopper Oxidase Engineering Database (LccED). Genome mining revealed the presence of three MCOs with putative catalytic activity – two MCOs homologous to bilirubin oxidases (designated SF1.4_MCO1 and SF1.4_MCO2), and a third MCO-like multicopper polyphenol oxidase (MPO) – designated 2-8_MPO - showing homology to YfiH/Domain of Unknown Function (DUF152) family of proteins. A rudimentary analysis of the characteristics of these MCOs were explored. The MCOs and MPO were synthesised (in pET-28a(+) vectors) and transformed into *Escherichia coli* BL21, and functional expression was confirmed through the oxidation of 1 mM ABTS (in 0.05 M sodium acetate, pH 3.0). Primary activity-pH and activity-temperature dependency was determined between pH 2.0 – 5.0 and a temperature range of 25˚C-45˚C.

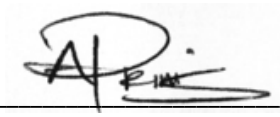
These are the first reports of MCOs from these South African environments, signifying that these environments may be a great source of undiscovered actinobacteria with biotechnological potential, especially as they were isolated from rare actinobacterial genera.

Keywords: Cape Floristic Region, Rooibos, Peat, Laccase, Multicopper Oxidases, MPOL

DECLARATION

I declare that "*Actinobacteria associated with two diverse soil environments and their multicopper oxidase diversity*" is my own work, that it has not been submitted for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated or acknowledged by complete references.

Full Name: _____ALARIC PRINS_____ Date: ____11 NOVEMBER 2023_____

Signed: 

ACKNOWLEDGEMENTS

- I would like to thank my supervisors, Prof. Le Roes-Hill and Prof. Kirby-McCullough for their immense support and guidance. Thank you for your patience and insights, and for providing me with the tools to grow as a scientist.

- To my parents, Mackie and Adam, for their love and support and words of encouragement. A special thank you to my mother, whose unwavering work ethic I've inherited. Thank you for teaching me the value of hard work in pursuing my dreams.

- To my brother and sister (in-law), Marlon and Bronwyn, my nephew and niece, Kevin and Kayla – my second home, you mean the world to me.

- My friends, Kim, Matthew, Cindy, Jude and Melissa – my eternal safe space. Thank you for the love, for the laughs, the cries, and the never-ending words of encouragement. Thank you for pushing me to do my best, always.

- The Boois - Kutala and Siya – your unwavering support and friendship to me and my family will always be remembered.

- Prof. Dr. Jürgen Pleiss and Mr. Max Haüssler – thank you for your collaboration, insightful views about enzymes and introducing me to the world of computational biology. I have a lot to learn! Here's to many more EnzymeML meetings.

- The members of BioAct, my friends and AMHBI lab colleagues at CPUT – thank you for providing a fun and challenging environment. I look forward to working together over the coming years and explore the wonderful world of actinobacteria.

Research Output(s)

2023
SASM 2023 – Oral Presentation
Prins, A., Hauessler, M., Kirby-McCullough, B., Pleiss, J. and Le Roes-Hill, M.
Accessing the biodiversity of two distinct South African soil environments: the search for novel oxidative enzymes
Stellenbosch, Western Cape, South Africa

4th EnzymeML Workshop – Oral Presentation
Prins, A., Hauessler, M., Le Roes-Hill, M.
EnzymeML – proof-of-concept study using SLAC to determine the kinetic parameters using a Python-based Jupyter notebook
Rüdesheim, Germany, organised by Dr. Carsten Kettner (Beilstein Institute) and Prof. Dr. Juergen Pleiss (IBTB, University of Stuttgart), 17-19 October 2022

2022
3rd EnzymeML Workshop
Rüdesheim, Germany, organised by Dr. Carsten Kettner (Beilstein Institute) and Prof. Dr. Juergen Pleiss (IBTB, University of Stuttgart), 17-19 October 2022

2021
CPUT Research Festival
26 October 2021
Agunbiade M., Ngobeni R., **Prins A.**, Le Roes-Hill M.
Application of actinobacterial oxidative enzymes

2020
EnzymeML Workshop – Systematic profiling of multicopper oxidases
5 – 7 October 2020, South Africa-Germany Bilateral Research Projects, organised by Dr. Marilize Le Roes-Hill (AMHBI, Cape Peninsula University of Technology) & Prof. Dr. Juergen Pleiss (IBTB, University of Stuttgart)

2019
FEMS 2019 – Poster Presentation
Glasgow, Scotland, United Kingdom
Prins A., Kirby-McCullough BM. and Le Roes-Hill M.
Actinobacteria associated with Rooibos and a South African peatland, and their multicopper oxidases

2019
Research Exchange – Fraunhofer Institute hosted by Prof. Dr. Volker Sieber
Straubing, Germany
Two-month NRF-funded Research Exchange training on cloning procedures in E. coli.

2018
BIOCAT 2018 – Poster Presentation
Hamburg, Germany
Prins, A., Kirby-McCullough, B. M. and Le Roes-Hill, M.
Actinobacteria associated with Rooibos and a South African peatland, and their multicopper oxidases

2017
14th Symposium on Bacterial Genetics and Ecology – Poster Presentation
Aberdeen, Scotland, United Kingdom
Prins, A., Kirby, B. M. and Le Roes-Hill, M.
Actinobacteria associated with Rooibos (Aspalathus linearis) and their multicopper oxidases

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

CHAPTER ONE
INTRODUCTION

1.1 Background

Multicopper oxidases (MCOs) are a large family of enzymes that includes bilirubin oxidases, ascorbate oxidases, ceruloplasmin, phenoxazinone synthases and the largest subfamily - laccases (Solomon *et al.*, 1996). Of these MCOs, laccases have been extensively exploited for their use in industrial processes. MCOs are able to oxidise substrates to generate reactive radicals which may be involved in a variety of reactions, including polymerisation through the oxidative coupling of monomers, degradation of polymers or the degradation of phenolic compounds (Claus, 2004). Additionally, in the case of steric hindrance, the oxidation of larger, more complex molecules is directly inhibited, but can be chemically mediated with smaller redox mediators (Roth & Spiess, 2015). Laccases have been extensively used in industrial oxidative processes including dye decolourisation, stain bleaching, bioremediation of environmental phenolic contaminants, pulp and paper processing, ethanol production, delignification, and plant fibre modifications (Alcalde, 2007).

The phylum *Actinomycetota* is one of the largest and most diverse groups within the domain Bacteria (Stackebrandt & Schumann, 2006). They are of a significant biotechnological importance, particularly the strains within the order *Actinomycetales,* as well as members of the orders *Kitasasporales, Pseudonocardiales* and *Streptosporangiales*, due to their ability to produce a large number of industrially relevant secondary metabolites, such as bioactive compounds and enzymes (Bérdy, 2005; Anandan *et al.,* 2016; Goodfellow *et al.*, 2018). Studies on the isolation and characterisation of MCOs from actinobacteria are rare (Alves *et al.*, 2014), especially from the “rare”, not-readily isolated non-streptomycete genera, and these taxa will likely serve as a potential source of novel MCOs.

1.2 Problem Identification

Fungal MCOs have been significantly exploited for their applications in industrial processes. More recently, however, MCO and MCO-like enzymes from bacteria have become increasingly prominent (Claus, 2004). This is facilitated by characteristics that not only mirror those of fungal MCOs, but also the possession of unique and industrially important traits, such as enhanced thermostability and the ability to oxidise substrates under more alkaline conditions as opposed to more acidic conditions required for fungal MCOs (Reiss *et al.*, 2011). Additionally, bacterial MCOs may prove advantageous for biotechnological development attributable to their atypical characteristics, shorter production times and ease of genetic manipulation allowing for better expression in heterologous systems (Burton, 2005; Santhanam *et al.*, 2011). The small laccase (SLAC) from *Streptomyces coelicolor*, for example, is a bacterial enzyme that was identified through a genome mining approach, which was subsequently cloned and characterised in several studies (Machczynski *et al.*, 2004; Skálová *et al.*, 2009). Despite this wide range of desirable characteristics, bacterial MCOs are shown to have a low redox potential (430 mV for SLAC) compared to a redox potential of up to 800 mV which has been reported for basidiomycete laccases, making them less powerful catalysts and reducing their commercial value. It is, therefore, necessary to search for laccases of bacterial origin that are equivalent to laccases of fungal origin.

The Cape Floristic Region (CFR) is one of the most biodiverse regions in South Africa boasting unique plant diversity with a rich concentration of endemic plants. Unfortunately, the CFR is experiencing drastic habitat loss (Stafford *et al.*, 2005). *Asphalathus linearis* (Rooibos) is an indigenous plant that grows in the southern regions of the Western Cape and is cultivated for the production of Rooibos (“red bush”) tea (Standley *et al.*, 2001). The beneficial properties of this herbal tea are, in part, attributed to a variety of phenolic compounds produced by the plant (Van Heerden *et al.*, 2003). Emerging peatlands in the CFR has garnered attention over recent years, through studies that aim to understand the microbial diversity and biogeochemical cycling that drives the formation of these unique environments. Limited research is available on the actinobacterial diversity associated with Rooibos plants and emerging peatlands, and information on potential enzyme production by these strains is virtually non-existent.

This holds particularly true for the MCO complement of the rhizospheric microorganisms from these environments. Considering the high phenolic constituents of the plant materials in these environments that may drive oxidative enzyme production, it is necessary to investigate the diversity of the actinobacteria associated with the plant and the unique environmental pressures, to determine whether it drives the expression of MCOs, whilst also gaining access to potentially novel bacterial MCOs that may possess similar or better activity profiles than that of fungal MCOs.

1.3 Aims and Objectives

- *Perform a rudimentary analysis of the actinobacterial diversity of the study sites at the time of sampling:*
    - through metataxonomic analyses to identify the major actinobacterial constituents to family and genus level using actinobacterial-specific 16S rRNA sequencing primers
    - determine the physicochemical parameters of the soil environment at the time of sampling and relate it to the actinobacterial diversity.
- *Isolate new actinobacterial strains from the selected study sites:*
    - through selective isolation techniques, with a particular emphasis on selecting strains that are typically underrepresented (non-streptomycetes)
    - screening the isolates for the production of extracellular MCOs through colorimetric assays of liquid culture supernatants
    - select the most talented strain from each environment for genome sequencing
    - perform a basic genome analysis of the selected strains to screen for novel MCO genes
    - perform a rudimentary analyses of protein features for a select number of newly identified MCOs
    - clone and determine whether heterologously expressed MCOs are functional
- *Develop a Python-based workflow for kinetic parameter estimation*
    - by designing a microtiter plate (MTP) workflow that takes into consideration a range of technical criteria

- Determine the kinetic parameters of the small laccase from *Streptomyces coelicolor* using a non-conventional Python-based workflow and comparing it to reported biochemical data in literature.

1.4 Variations to original study design

At the inception of the study, only the Rooibos environment was selected as a study site. However, a second site (peat from Springfield Estate, Cape Agulhas) was introduced a year after the inception of the project, when a concurrent project found that a rare actinobacterium isolated from this environment harboured a novel MCO gene. Therefore, it was decided that this novel gene should be explored further in this study, along with a selected isolate from the Rooibos site.

CHAPTER TWO

LITERATURE REVIEW

2.1 Cape Floristic Region

Influenced by its distinct climate, South Africa is divided into nine terrestrial biomes, viz. Succulent Karoo, Nama Karoo, Fynbos, Forest, Albany Thicket, Grassland, Savanna, Indian Ocean Coastal Belt and Desert (Figure 2.1; Rutherford *et al.*, 2006). Despite only covering approximately 90 000 km^2 of landmass, the Cape Floristic Region (CFR) is one of the most prolific South African biomes, containing over 8 500 plant species (Meadows & Sugden, 1993). Of these plant species, almost 75% are endemic to the region (Goldblatt, 1978) and the CFR has been identified as a global biodiversity hotspot (Myers *et al.*, 2000). To further emphasise this high endemism, a comparative study by Born *et al.* (2007) of 11 530 native seed plant species across five biome regions, showed that 8 856 are found in the CFR when compared to the other four regions.

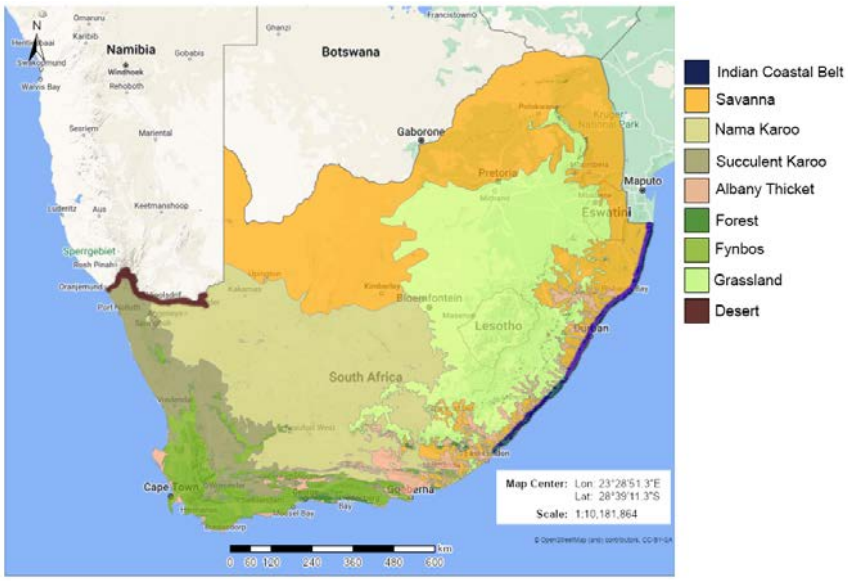


Figure 2.1: Map of South Africa outlining the nine biomes described by (Low & Rebelo, 1998); map generated using Cape Farmer Mapper 3 using Low and Rebelo (2006) overlay, with amendments to include the Desert and Indian Coastal Belt

2.1.1 Fynbos

The Fynbos ("*fine bush*") biome, which forms part of the Greater CFR, can be characterised by the prevalence of low to medium height shrubland, including 81 distinct true and fire-prone fynbos vegetation units, as well 38 sandveld and renosterveld units. The Fynbos biome is located in the Western Cape, South Africa and starting from the Boland it extends in a westerly and northward direction from the Cape Peninsula to Vanrhynsdorp, and easterly towards Makhanda (formerly Grahamstown) (Low & Rebelo, 1998).

Fynbos species are typically found in soils that are acidic in nature (pH 4.0-7.0), and are relatively nutrient-poor (Manning, 2018). The occurrence of high levels of edaphic endemism is evidence of this – studies performed by Stock and Midgley (1995) and Stewart *et al.*, (1993) surveying nitrogen utilisation in fynbos and Australian kwongan (a fire-prone woodland similar to fynbos), respectively, highlighted important differences in soil nutrient availability in determining the spatial distribution of plant species in nutrient-poor environments.

Furthermore, such vegetation-environment relationships have been determined in the Soetanysberg, Cape Agulhas (Richards *et al.*, 1995). The Agulhas plain, which also forms part of the CFR, is a coastal lowland that covers approximately 1 600 km^2 of landmass at the southern tip of Africa (Figure 2.2). It is characterised by transgression events during the Miocene and early-to-mid Pliocene eras (Hendey, 1983). It is a geologically diverse region, with most of the associated soil types (deep acid sands, shallow acid sands and shallow alkaline soils) postdating the regression events (Hendey, 1983).

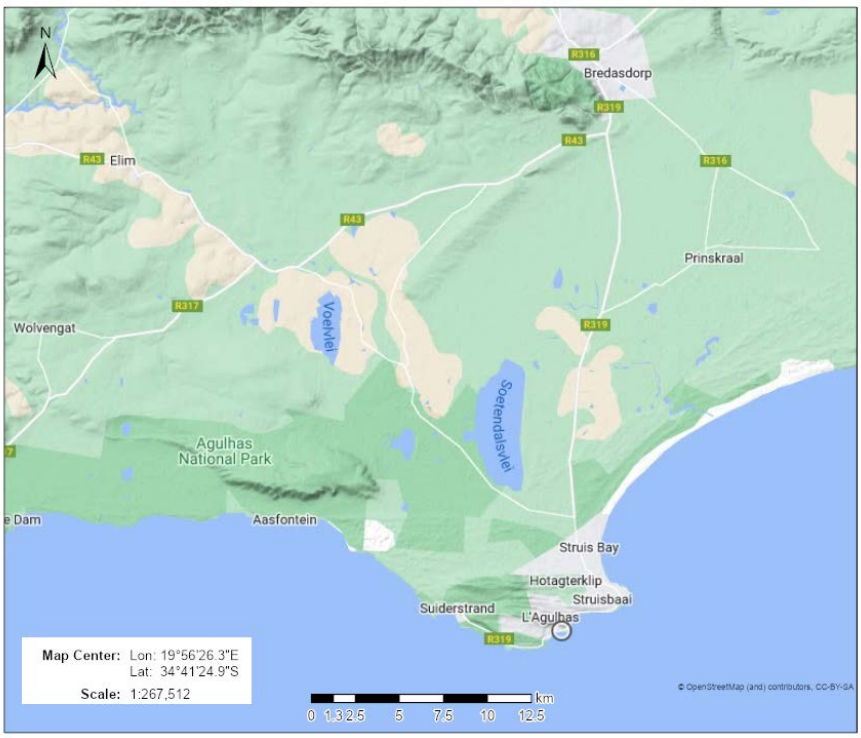


Figure 2.2: Overview map of a section of the Agulhas Plain; map generated using Cape Farm Mapper 3

In alignment with typical fynbos soil types, Richards *et al.* (1997) demonstrated that nitrogen and phosphorus availability in the soil matrix is an important factor in landscape-level species distribution in nutrient-poor Mediterranean-climate ecosystem, which is supported by previous studies that revealed high species turnover across soil types that differ in pH and nutrient content (Cowling, 1990; Thwaites & Cowling, 1988).

2.1.2 Peat

One of the key points of interest in the Agulhas Plain is the presence of peatlands. Peat is formed by the gradual decay of plant material (including moss, grass, shrubs, and trees) under sustained waterlogged conditions which results in the formation of brown-black organic-rich soil (Ellery *et al.*, 2012). Peat is widespread, globally, with an estimated coverage of 4.2 million km^2 of terrestrial and freshwater surface – although it is predominantly found in temperate-cold climates of the Northern Hemisphere (Ivanova *et al.*, 2020).

The occurrence of peat and peatlands in South Africa are relatively rare. Their presence in South Africa is typically associated with coastal plains where the regional groundwater table intersect with depressions, or in catchment areas of quartzite and dolomite (Job & Ellery, 2013). Grundling *et al.* (2017) identified sixteen peatland eco-regions in South Africa and calculated their distribution based on combining scoring models from 2011 and 2016 (Table 2.1).

Table 2.1: Table representing known peatland points scored based on the combined peatland ecoregions per the 2016 model sorted in descending total percentage (adapted from Grundling *et al.*, 2017).

Peatland Ecoregion	Count	Percentage
Natal Coastal Plain	343	63.1
Central Highlands	82	15.1
Highveld	38	7.0
Great Escarpment Mountains	31	5.7
Lowveld	20	3.7
Southern Coastal Belt	20	3.7
Cape Folded Mountain	8	1.5
Eastern Coastal Belt	8	1.5
Bushveld Basin	2	0.4
Eastern Uplands	1	0.2
Limpopo Plain	1	0.2
Ghaap Plateau	0	0
Great Karoo	0	0
Nama Karoo	0	0
Southern Kalahari	0	0
Western Coastal Belt	0	0
Total Model Points	*554*	*100*

Peat-forming systems are ecologically significant as they collectively store up to 500 – 700 billion tonnes of carbon, more than what is stored in the world's tropical rainforests and equal to that of atmospheric carbon (Grace, 2004; Pan *et al.*, 2011; Parish *et al.*, 2008; Yu *et al.*, 2010). Peatlands can be considered as a potential driver of climate change due to it being a prospective source of carbon emissions, particularly if severely impacted as a result of human activities and/or natural disasters (Frolking *et al.*, 2006; Gründling & Grobler, 2005; Job & Ellery, 2013).

Particularly, within the South African ecosystem context, these disturbances include erosion (which includes anthropogenic activities such as farming), flooding disproportionate to rainfall and desiccation of the landscape which leads to biodiversity losses and release of greenhouse gasses (Job & Ellery, 2013). These disruptions are alarming, since 65% of South African wetlands are in a critical condition, and more than 50% have been destroyed (Cowan, 1995; Driver *et al.*, 2012).

As such, concerted efforts are geared towards conservation and regeneration of peatlands. For example, Rebelo *et al.* (2019), as part of a comparative study, identified three pristine palmiet wetlands (*Pronium serratum*) in the CFR (Theewaterskloof, Goukou, and Kromme) that are currently assigned for private agricultural use. They identified that these palmiet wetlands sequester 21-41g.m^{-2} of carbon per year, have nitrogen and phosphorus uptakes of 62-85% and reduces flooding 16 times more than when compared to degraded wetlands. Their findings emphasise the importance of policies between landowners and cities to ensure the best-use case for valuable ecosystems.

2.1.3 Rooibos

Rooibos (*Aspalathus linearis*, literal: "*red bush"*) is an endemic South African fynbos species (Figure 2.3). It is a seed plant that is world-renowned as an herbal tea. The genus *Aspalathus* is made up of more than 270 species, which are predominantly endemic to the CFR (Dahlgren, 1968).

Aspalathus linearis (legume, family: *Fabaceae)* is a polymorphic plant, with needle-like leaves, with distribution mainly along the Cederberg Mountain region of the Western Cape (Figure 2.4). *Aspalathus linearis* mainly grows in acidic soils (pH 3-5.3) that are well-drained, nutrient poor and are derived from sandstone (Muofhe & Dakora, 2000). Cultivation of *A. linearis* beyond its native range has proven to be a challenge, and the cause of this endemicity has yet to be determined.

Figure 2.3: An image of a flowering Rooibos (*Aspalathus linearis*) plant taken in Clanwilliam, Western Cape (image taken by Winfried Bruenken, CC BY-SA2.5)

Rooibos is a commercially important plant. Historically, the San and Khoi indigenous people of Southern Africa were the first to utilise Rooibos as an herbal tea (DEA, 2014). Per annum, approximately 20 000 tons of Rooibos is produced with annual incomes of ca. 500 million ZAR (Schroeder *et al.*, 2020; Wynberg, 2023). In terms of market share, Rooibos accounts for 10% of the global herbal tea exports to more than 30 countries, with Rooibos Limited contributing up to 70% of the total Rooibos production (Marie-Vivien & Biénabe, 2017; Troskie & Biénabe, 2013).

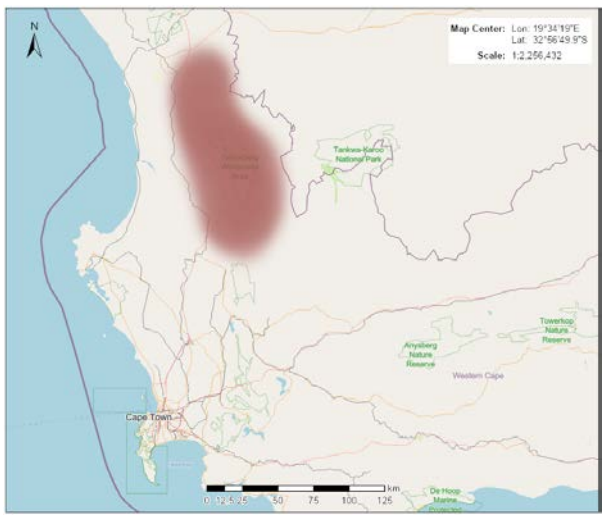


Figure 2.4: Distribution (red) of native Rooibos (*Aspalathus linearis*) cultivation in the Cederberg Mountain area of the Western Cape (map generated with Cape Farm Mapper 3, amended to highlight distribution area)

The chemical composition of extracts from Rooibos has been extensively studied. The major phenolic constituents present can be divided into two classes: glycosides, including aspalathin (Mw: 452 g.mol^{-1}) and isoorientin (Mw: 448 g.mol^{-1}), and the aglycones, including quercetin (Mw: 302 g.mol^{-1}) and luteolin (Mw: 286 g.mol^{-1}) (Joubert *et al.*, 2009; Figure 2.5).

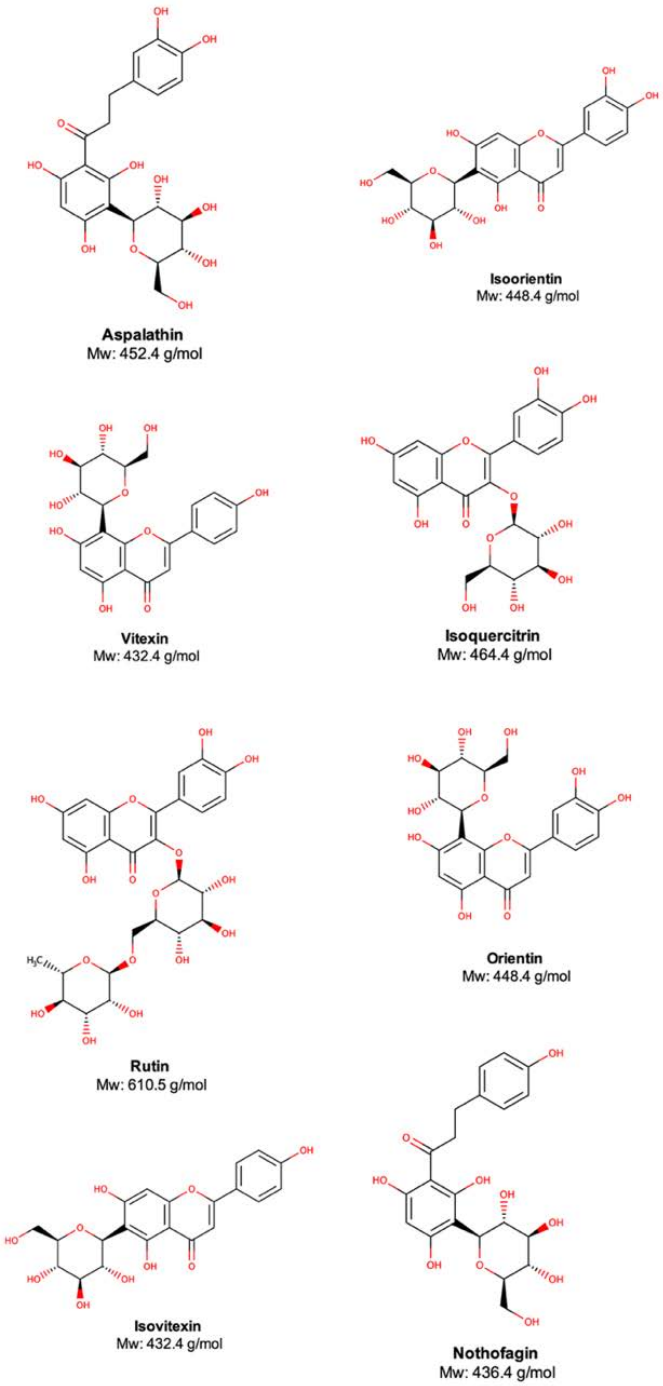


Figure 2.5: Major phenolic constituents in Rooibos (adapted from Joubert *et al.*, 2009); chemical structures drawn in Marvin Sketch version 23.12)

From the scientific development of indigenous ethnomedicine, various studies have explored the health benefits of the various phytochemicals from Rooibos. Among these benefits are chemoprotective effects (Huang *et al.*, 2019; Magcwebeba *et al.*, 2016; Marnewick *et al.*, 2009), protection against neurodegeneration (Inanami *et al.*, 1995; Minné *et al.*, 2023) and other antioxidant activities which includes protection against cardiovascular disease (Fekry *et al.*, 2014; Marnewick *et al.*, 2011; Pantsi *et al.*, 2011), anti-inflammatory effects (Baba *et al.*, 2009; Pretorius & Smith, 2022) and anti-mutagenic properties (Erickson, 2003).

2.1.4 Actinobacterial diversity associated with Rooibos and peat soils

Microorganisms are key role-players in biogeochemical cycles through a range of plant-soil microbe interactions that are important in maintaining plant productivity and soil ecosystems (Le Roux *et al.*, 2017). Disturbances (particularly of anthropogenic origin) to optimal environmental conditions, especially where unique conditions are required, can therefore lead to a reduction in plant persistence (Béna *et al.*, 2005). In the case of legumes from the family *Fabaceae*, which *A. linearis* is classified under, rhizobia form root nodules that aid in nitrogen-fixation to provide the host plant with organic forms of nitrogen, which is exchanged for nutrients from the plants to the microbes (Franche *et al.*, 2009). This is performed in tandem with other free-living soil microorganisms that are involved in carbon cycling, maintaining the soil structure and biological control (Kennedy, 1999).

Several studies have been performed to assess the influence of the soil microbial diversity on the growth of Rooibos. Hassen *et al.* (2018) performed a culture-based study to determine the colonisation frequency of nodule and endophytic bacteria associated with healthy and declined Rooibos plants. More than 75% of isolates from healthy plants belonged to the *Rhizobium-Bradyrhizobium* group with the rest identified as *Burkholderia*, *Pseudomonas* and *Bacillus* species. Interestingly, a low number of root-nodulating rhizobia were detected in the plants with declined health. The authors hypothesise that external factors such as soil acidification and nutrient depletion could prevent migrating free-living organisms from forming nodules prior to symbiosis, leading to unfavourable conditions for plant survival. A similar effect was seen in a previous study

by the same authors (Ahmed *et al.*, 2014) where nodulation of soybean failed as a result of limiting factors.

Brink *et al.* (2020) examined the bacterial communities associated with natural and commercial Rooibos plants during the dry and wet season through terminal restriction fragment length polymorphism (t-RFLP) analyses and real-time polymerase chain reaction (RT-PCR). The study showed that *Actinobacteria*, *Proteobacteria* and *Acidobacteria* species were the most dominant bacterial phyla detected, with large similarities between both environments. The relationship between physicochemical parameters and the microbial compliment was also assessed. Members belonging to the family *Pseudomonadaceae*, an indicator taxon, showed a strong positive correlation with Na^+ and K^+ levels during the dry season, whereas the *Bradyrhizobiaceae* were an indicator for the wet season, correlated negatively. Of note, and in accordance with the more recent study by Brink *et al.* (2020), Postma (2016) also found that while microbial diversity differed by season (which is likely linked to rainfall and the associated shifts in environmental conditions due to run-off), the overall classification were dominated by *Acidobacteriales* and *Actinomycetales*, suggesting these organisms play an important role in these low nutrient environments that is characteristic of fynbos.

The microbial diversity associated with peat is well documented, from a range of different environments ranging from boreal forests in Finland, peat-bound Karstic wetlands in Australia containing stromalites, arctic soils, as well developing peatlands on the Agulhas Plain (Makhalanyane *et al.*, 2016; Proemse *et al.*, 2017; Sun *et al.*, 2014; Weels *et al.*, 2022).

Whilst fungi are considered to be the main drivers of decomposition in the carbon-rich environments (with a wide array of lignocellulosic enzymes to breakdown recalcitrant materials), actinobacteria have been shown to exhibit similar enzyme activities that also play a large role in mineralisation processes (Kusai *et al.*, 2018; Lew *et al.*, 2018; B. Liu *et al.*, 2020). In fact, across a number of peat diversity studies, actinobacteria contribute a significant proportion of the prokaryotic complement. For example, in boreal peat samples, actinobacteria had the highest relative abundance of 24.5% (Sun *et al.*, 2014). Similarly, in a study that sampled across three wetlands systems, for three different

sampling depths ranging from 0-50cm, actinobacteria had the highest relative abundance across all samples (Weels *et al.*, 2022). These organisms have the ability to utilise a variety of different substrates and easily adapt to a range of temperatures, as evident in a study where actinobacterial diversity was detected up to depths of 110cm In a Siberian permafrost peatland (Aksenov *et al.,* 2021).

2.2 Actinobacteria

Actinobacteria are Gram-positive bacteria (Figure 2.6) that are mostly free-living and widely distributed in terrestrial and aquatic (including marine) environments (Macagnan *et al.*, 2006). They typically have large genomes, with a high %mol G+C content of their genomic DNA.

Actinobacteria exhibit a multitude of different cell morphologies, ranging from rod (*Mycobacterium*), fragmenting-hyphae (*Nocardia*), spore-bearing branched-hyphae (*Micromonospora*) to branched mycelia (*Streptomyces*) and mycelia-free (*Corynebacterium*) (Barka *et al.*, 2016).

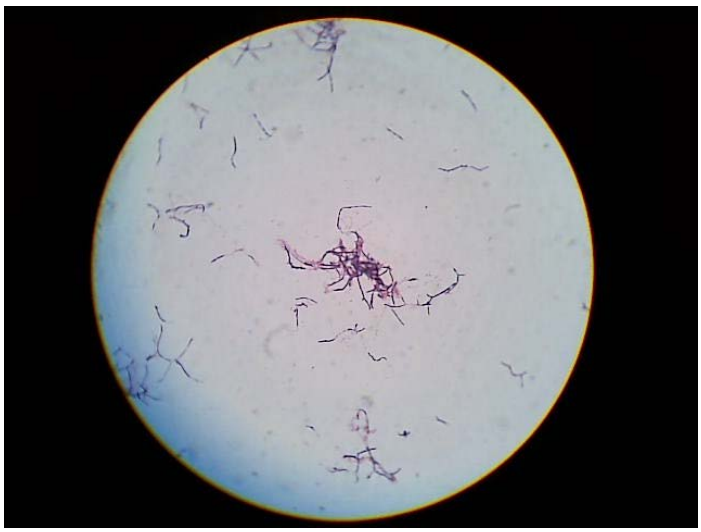


Figure 2.6: A microscope image of an actinobacterium (*Micromonospora* sp.) showing distinct filamentous structure (own image).

Actinobacteria possess extraordinary properties, which is essentially attributed to their large genomes and complex morphological changes in their lifecycles. As such, they are often able to withstand a number of environmental pressures such as extremes of temperature (hot or cold), varying chemical conditions (acidity, alkalinity, high salinity, nutrient-poor environments and moisture) (Goodfellow *et al.*, 2018; Mohammadipanah & Wink, 2016; Shivlata & Satyanarayana, 2015). This adaptability allows them to thrive in a wide range of extreme and special environments, including but not limited to Arctic and Antarctic regions, glaciers, deserts, caves, hot springs, and mangroves (Law *et al.*, 2020).

The phylum *Actinomycetota* Goodfellow, 2021 (previously *Actinobacteria*) (Oren & Garrity, 2021) is one of the largest phyla in the domain Bacteria (Stackebrandt & Schumann, 2006). Currently, the phylum is divided into six major classes, namely *Acidimicrobiia, Actinomycetes, Coriobacteriia, Nitriliruptoria, Rubrobacteria* and *Thermoleophilia* (LPSN, DSMZ, https://www.bacterio.net/, date accessed 28 October 2023). These can be further divided into a total of 22 orders, 54 families and 250 genera (Ludwig, Euzéby, & Whitman 2012; Table 2.2).

Table 2.2: A summary of the total number of organisms in the phylum *Actinomycetota* (table adapted from Ludwig *et al.,* 2012)

Phylum	Class	No. of Orders	No. of Genera
Actinomycetota	I – *Actinomycetes*	15	249
	II – *Acidimicrobiia*	1	5
	III – *Coriobacteriia*	1	13
	IV – *Nitriliruptoria*	2	2
	V – *Rubrobacteria*	1	1
	VI – *Thermoleophilia*	2	4

Perhaps the most renowned of these are the *Streptomyces* (phylum *Actinomyetota*; class *Actinomycetia*; order *Streptomycetales;* family *Streptomycetaceae*) which have been extensively studied due to their prolific production of antimicrobial compounds (Lee *et al.*, 2018). Beyond their ability to produce bioactive metabolites (including ones with anti-tumour and anti-inflammatory properties), they are also important in their role as

plant-associated symbionts as well as the production of industrially relevant enzymes (Barka *et al.*, 2016; Le Roes-Hill & Meyers, 2009; Qin *et al.*, 2016; Silva *et al.*, 2020; Su *et al.*, 2021).

2.3 Multicopper Oxidases

2.3.1 Oxidative enzymes from actinobacteria

There is an increased interest in the production of oxidative enzymes from actinobacteria, particularly when it comes to applications involving the breakdown of lignocellulosic materials and detoxification of environmental pollutants including polycyclic aromatic hydrocarbons (PAH), micropollutants such as microplastics and endocrine disruptors (e.g., bisphenol-A), organophosphate pesticides and azo dyes (Torres *et al.*, 2003; Le Roes-Hill & Prins 2016).

As of the February 2023 update of BRENDA (a comprehensive enzyme database) more than 3 600 records match oxidoreductase production by actinobacteria, therefore within the context of this study, only multicopper oxidases will be discussed, with specific reference to laccases (EC 1.10.3.2), bilirubin oxidases (EC 1.3.3.5) and laccase-like copper oxidases.

2.3.2 The primary, secondary and tertiary structure of MCOs

Multicopper oxidases (MCOs) are a group of enzymes that contain between one to six copper atoms per molecule. Oxidation of a substrate occurs at a mononuclear T1 copper centre (Figure 2.7), after which electrons are shuffled to the T2 (1x Cu^{+}) and T3 (2x Cu^{+}) trinuclear cluster (TNC) where the reduction of dioxygen occurs, yielding two water molecules (Giardina *et al.*, 2010).

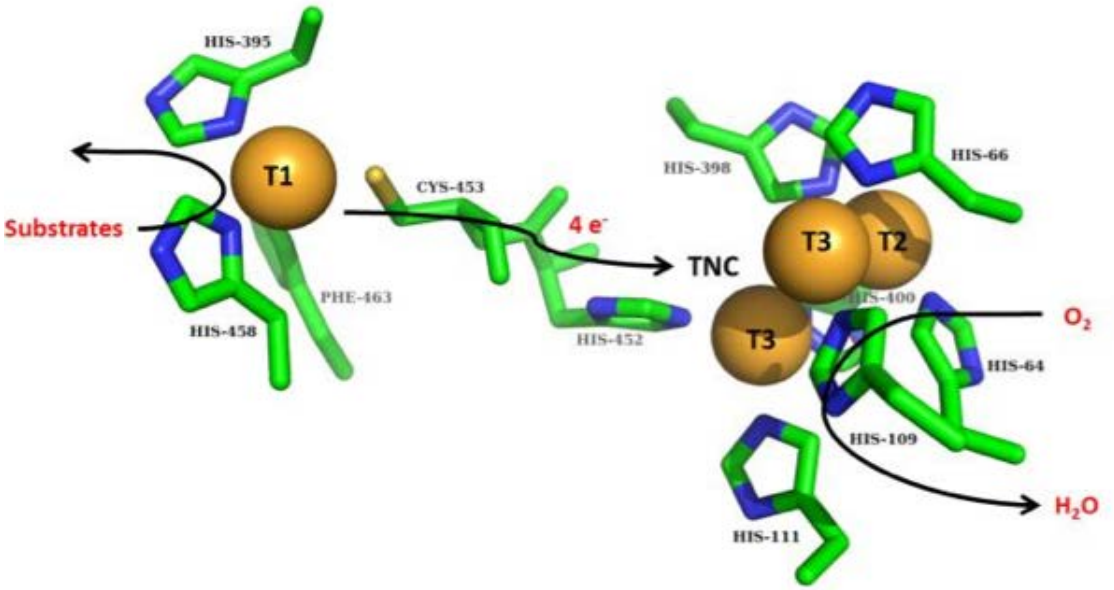


Figure 2.7: 3D model depicting the typical flow of electrons from the T1 copper to the TNC in the active site of a multicopper oxidase during the oxidation of a substrate (Prins *et al.,* 2015)

A highly conserved amino acid signature sequence is typically associated with the copper binding site of all MCOs: M2 signature pattern = G-X-[FYW]-X-[LIVMFYW]-X-[CST]-X-{PR}-{k}-X2-{S}-X-{LFH}-X3-[LIVMFYW]; M4 signature pattern = H-C-H-X3-H-X3-[AG]-[LM] (PROSITE entry number - PDOC00076; http://prosite.expasy.org) (Messerschmidt & Huber, 1990; Sigrist *et al.*, 2009).

Two conserved cysteine residues and a histidine are trigonally coordinated to the T1 copper forming a metalloorganic bond, with a fourth variable axial ligand – typically a methionine in bacterial MCOs. The T2 copper is coordinated with two histidine residues, while the two T3 coppers are coordinated to a total of six histidine residues (Dwivedi *et al.*, 2011; Enguita *et al.*, 2003).

In terms of overall structure, MCOs typically have two, three or six cupredoxin-like domains (Figure 2.8), arranged in Greek-key beta-barrel made up of two β-sheets with four strands each (Giardina *et al.*, 2010; Nakamura *et al.*, 2003). This arrangement is essential for MCOs to maintain their functionality (Herrera-Zúñiga *et al.*, 2019).

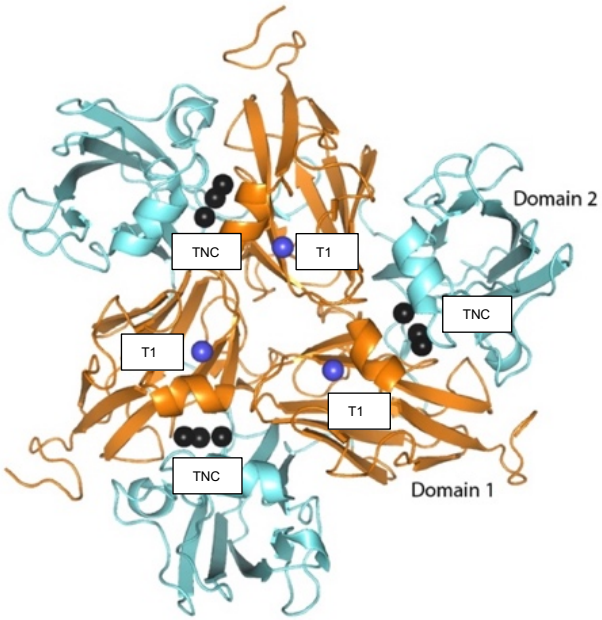


Figure 2.8: 3D structure of a two-domain MCO [*Streptomyces coelicolor* small laccase: PDB ID: 3CG8, (Prins, 2015)]

2.3.3 Reactivity and uses of MCOs

While the majority of commercially exploited fungal laccases operate at an optimal pH range of 3.5-5.0, bacterial laccases have been shown to exhibit oxidation of substrates at much higher pH with pH optima up to pH 9.0 (Morozova *et al.*, 2007; Reiss *et al.*, 2011). Redox potential at the active site (which is a measure of the enzyme's ability to extract electrons from substrate) typically ranges from 200-800mV, with the redox potential of bacterial MCOs often limited to the lower range (Moreno *et al.*, 2020).

However, given the ease of manipulation of bacterial genes, mutagenesis can be useful to modulate residues surrounding the active site to increase redox potential. For example, in a study to increase the redox potential of the two-domain small laccase from *Streptomyces coelicolor* (SLAC), several single-site mutations to the axial ligand and surrounding residues were performed. The best mutation was acquired by changing a valine at position 290 to an asparagine, with an observable increase in specific activity of

5× the wild-type SLAC, a shift in pH optima and an effective increase in redox potential from 365mV to 485mV (Prins *et al.,* 2015; Prins *et al.* unpublished data).

MCOs are extremely diverse in their substrate utilisation. They can oxidise a wide variety of compounds, from aromatic amines and thiols, substituted phenols, lignin rich aromatic compounds to siderophores and pigments, and as such, they are often applied in a number of industrial applications (Alcalde, 2007; Sharma & Kuhad, 2009; Dwivedi *et al.,* 2011). Four substrates typically used in the characterisation of MCOs are shown in Figure 2.9.

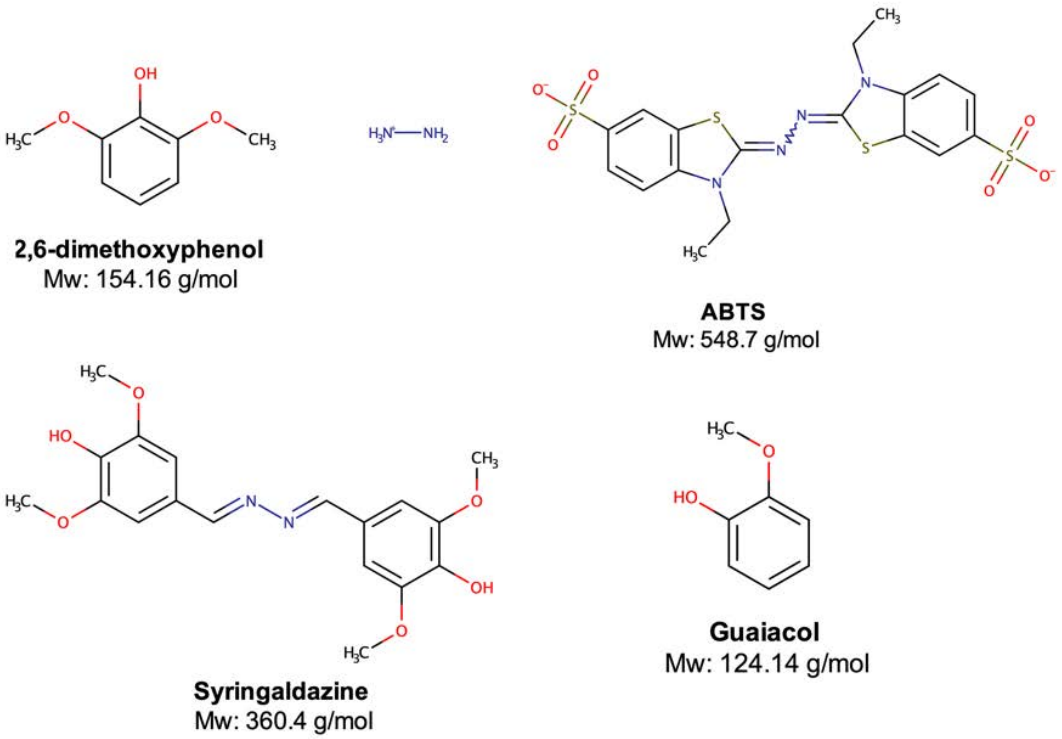


Figure 2.9: Commonly used substrates to measure MCO activity spectrophotometrically (chemical structures drawn in Marvin Sketch 23.12)

Laccases (EC 1.10.3.2; LAC; benzenediol:oxygen oxidoreductase) are able to oxidise a wide range of phenolic and non-phenolic substrates and have mainly been applied in the delignification and modification of plant fibres, decolourisation of textile dyes, food processes such as the clarification of wine, as well as the removal of environmental contaminants (Alcalde, 2007; Burton, 2005; Cabana *et al.*, 2009; Rodríguez-Couto & Toca Herrera, 2006).

Bilirubin oxidases (EC 1.3.3.5; BOD; bilirubin: dioxygen oxidoreductase) oxidises bilirubin to biliverdin but, considering their structural and reactive similarities to other MCOs, it also has a similar substrate profile to laccases – oxidising common laccase substrate such as 2,6-dimethoxyphenol, syringaldazine, ABTS and guaiacol (Figure 2.9). Bilirubin oxidases are often applied in clinical applications for the detection of bilirubin in human serum, as electrochemical catalysts in biofuel cells and dye colourisation, natively or as immobilised in biocathodes (Liu *et al.*, 2009; Mano, 2012; Roucher *et al.*, 2019; Sakurai & Kataoka, 2007).

2.4 Discovery of MCOs

2.4.1 Culture-based screening

While the actinobacteria can produce a host of novel biocatalysts for industrial applications, enzymes can often be difficult to access from a purely culture-based perspective. Based on records available in BRENDA, the most reported oxidative enzymes from actinobacteria originate from isolated strains but a mere fraction of these are from strains belonging to the order *Actinomycetales* (Le Roes-Hill & Prins, 2016). "Rare" actinobacteria (non-*Streptomyces* species) are strains of actinobacteria that are considered less culturable than *Streptomyces* (Seong *et al.*, 2001) and, as such, are considered a good source of novel enzymes because they are largely unexplored (Suriya *et al.*, 2016). Several different isolation strategies can be employed to selectively isolate non-streptomycetes (Table 2.3).

Table 2.3: Strategies for the selective isolation of rare actinobacteria (adapted from Suriya *et al.* 2016)

Strategy	Benefit
Adding different antibiotics	Enhance selection of *Actinomycetales*[1]
Adding chemoattractants	Selection of *Actinoplanes*[2]
Super-high frequency radiation (SHF)	*Rhodococcus, Streptosporangium*[3]
Extremely high frequency (EHF), Ultraviolet (UV) radiation	*Nocardiopsis, Nocardia, Streptosporangium*[4]
Chloramine treatment	*Microtetraspora, Herbidospora, Streptosporangium, Microbispora*[5]
Buffering agents (e.g. $CaCO_3$), seawater	*Salinispora*[6]

[1] Hong *et al.* 2009; [2] Zhang & Zhang 2011; [3] Hayakawa 2008; [4] Bredholdt *et al.* 2007; [5] Hong *et al.* 2009); [6] Maldonado *et al.* 2005

Extracellular MCO activity from isolated strains can be screened for through the inclusion of indicator compounds (typically a common substrate) within agar media or within culture supernatants though colorimetric detection (Figure 2.10).

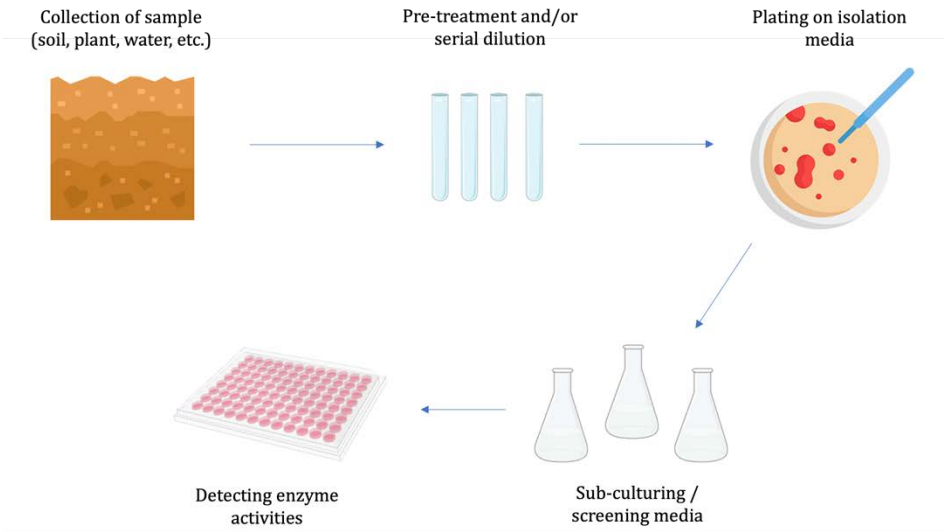


Figure 2.10: Simplified workflow of the isolation of actinobacteria for the detection of extracellular enzymes through liquid culture-based screening (diagram generated with icons from Biocons and Freeicons, CC-BY 3.0)

2.4.2 Sequence-guided mining of MCOs

It is believed that more than 99% of bacteria found in the environment are unculturable, leading to a necessity in exploring sequence space to access genes of potential interest (Locey *et al.*, 2017). Advances in next-generation sequencing has resulted in a marked decrease in the cost of genome and metagenome sequencing, and has given rise to an

exponential increase in the amount of publicly available sequence data (Zhang *et al.,* 2020). In fact, as of October 2023 more than 560 000 prokaryotic genome sequences were available on the National Centre for Biotechnology Information (NCBI) public database (https://www.ncbi.nlm.nih.gov/genome/browse#!/overview/; verified 28 October 2023). Similarly, more than 4.2 million publicly available BioSample resources (sequenced from various human and environmental sources) are available.

As such, unexplored sequence data (be it deposited or newly sequenced data) holds potential information on novel genes, particularly biocatalysts. These genes can be explored for their novelty and ideally expressed in a heterologous expression system to verify their biochemical capabilities (Figure 2.11). Various cases of successfully identified and cloned oxidases from actinobacteria have been reported including laccases from *Streptomyces sviceus* and *Streptomyces coelicolor*, a cholesterol oxidase (CO) from *Mycobacterium neoaurumi* and a FAD-containing Baeyer-Villiger monooxygenase (BVMO) from *Thermobifida fusca* (De Gonzalo *et al.*, 2012; Gunne & Urlacher, 2012; Machczynski *et al.*, 2004; Yao *et al.*, 2013).

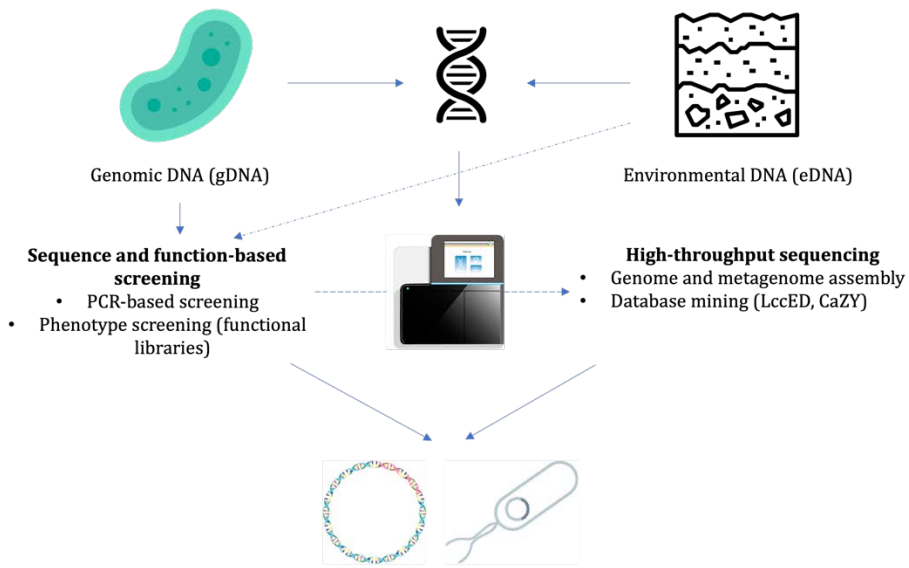


Figure 2.11: Simplified workflow for discovery of enzyme genes through genomic and metagenomic mining (diagram generated with icons from Biocons and Freeicons, CC-BY 3.0)

An interesting case of an unusual copper oxidase, termed a multicopper polyphenol oxidase (MPO) was discovered through a metagenomic mining study from bovine rumen microflora (Beloqui *et al.*, 2006). This four-copper oxidase, despite not harbouring the signature sequence of typical MCOs, exhibited activity against a wide number of "typical" laccase substrates including syringaldazine, 2,6-dimethoxyphenol, guaiacol and ABTS across a broad pH range. Interestingly, the redox potential at the T1 copper site was determined to be >700mV, much more comparable to that of typical fungal laccases. Since then, many other MPOLs of bacterial origin have been identified from metagenome screening studies performed on hot springs, compost, polluted soils, and lakes (Chai *et al.*, 2017; Jeon & Park, 2020; Narnoliya *et al.*, 2019; Sharma *et al.*, 2019).

Bioinformatics tools, with an emphasis on databases related to enzyme annotation and biochemical characterisation, are imperative for the successful mining of MCO genes from sequencing sources to provide sequence, structural and biochemical data which can be used to determine protein novelty. Primary protein sequence annotations can be performed on genome and metagenome sequences using databases such as the Rapid Annotation using Subsystem Technology (RAST) and UniProt (Aziz *et al.*, 2008; Bateman *et al.*, 2023). Well curated MCO-specific databases are especially useful as a rapid information source on sequence, structural and biochemical parameters. One such database is the Laccase and Multicopper Oxidase Engineering Database (LcEED). This database enables a localised BLAST environment and contains information on a total of 51 058 sequences and 229 elucidated structures assigned to 16 superfamilies and 105 homologous families (Table 2.4) based on sequence similarity and the occurrence of signature sequence motifs (Gräff *et al.*, 2020; Kumar *et al.*, 2003; Sirim *et al.*, 2011).

Table 2.4: An overview of the total number of MCO genes available in the Laccase and Multicopper Oxidase Engineering Database (LccED) – as verified on 28 October 2023

#	superfamily	Group	Homologous families	Proteins	Sequences	Structures
1	A - Basidiomycete Laccase	3dMCO	7	2034	2463	62
2	B - Ascomycete MCO	3dMCO	5	1584	1905	20
3	C - Insect Laccase	3dMCO	17	1096	1249	0
4	D - Fungal Pigment MCO	3dMCO	5	691	816	0
5	E - Fungal Ferroxidase	3dMCO	5	1144	1511	2
6	F - Fungal and Plant AO	3dMCO	7	2796	3415	4
7	G - Plant Laccase	3dMCO	6	3347	4024	0
8	H - Bacterial CopA	3dMCO	6	4594	7068	0
9	I - Bacterial Bilirubin Oxidase	3dMCO	12	3157	4315	30
10	J - Bacterial CueO	3dMCO	14	4859	9998	71
12	L - Bacterial MCO	3dMCO	7	5028	7951	2
11	K - SLAC-like (type B 2dMCO)	2dMCO	2	531	729	22
13	M - Archaeal type A 2dMCO	2dMCO	1	128	162	0
14	N - Bacterial type B 2dMCO	2dMCO	5	1671	2594	0
15	O - Archaeal and Bacterial type C 2dMCO	2dMCO	5	957	1265	9
16	P - Ceruloplasmin	6dMCO	1	1088	1593	7

2.4.3 Reproducibility of biochemical data

One of the largest challenges in enzyme characterisation is the lack of reliable and reproducible experimental data (Begley & Ioannidis, 2015). Studies have shown that there is a considerable discrepancy in the ability of researchers to reproduce their own results, as well as the findings of others as a result of the omission of critical information regarding experimental procedures. With regards to enzyme characterisation, this missing information often includes simple parameters such as enzyme and/or substrate concentrations, buffer compositions, pH values and assay temperatures (Baker & Penny, 2016; Halling *et al.*, 2018). In 2014, the Standards for Reporting of Enzymology Data (STRENDA) commission, published a set of guidelines to assist researchers in the thorough reporting of enzymology data including but not limited to protein information, assay conditions, detailed experimental methods and as well as the processing of results (Tipton *et al.*, 2014).

In addition to accurate reporting of biochemical data, data reusability is also an important factor for reproducibility. The generation of large datasets require depositing in accessible data repositories and there are additional requirements to ensure deposited data adheres to conditions for reusability and reproducibility, particular with respect to machine learning and computational biology tools. The FAIR (Findable, Accessible, Interoperable and Reusable) guidelines were published to ensure (i) data sources include metadata and conclusive vocabulary; (ii) that the data is accessible in an open and standardized protocol and is denoted with unique identifiers; (iii) it contains qualitative references, and (iv) that accurate and relevant attributes are extensively reported (Wilkinson *et al.*, 2016).

A number of studies have been performed that attempt to address gaps in literature and databases where critical experimental parameters are lacking through the generation of comprehensive datasets, building database tools that specifically curate biochemical data, or tools to assist in the curation and generation of biochemical datasets, all of which are in accordance with the STRENDA and FAIR guidelines. SABIO-RK, for example, is a biochemical user database that extracts data about biochemical reactions and enzyme kinetics from literature and makes it accessible with an application programme interface (API) for integration into other databases and biology tools and workflows (Wittig *et al.*, 2018), such as EnzymeML.

EnzymeML is an open data-exchange format that was developed by Range *et al.* (2022) to facilitate the capturing of biochemical data according to FAIR and STRENDA guidelines, and is incorporable into graphical user interfaces (BioCatHub, Malzacher *et al.*, 2020) or in Python-based Jupyter notebooks. Lauterbach *et al.* (2023) demonstrated the capabilities of EnzymeML through a multi-partner study that captured reaction data in the EnzymeML format, and through an EnzymeML toolbox, was able to incorporate the data into multiple modelling workflows. These included PySCes, COPASI and interferENZY – tools which allow for a variety of analyses including structural and stoichiometric analyses, parameter estimation and scanning, simulating biochemical networks, sensitivity analyses and determination of hidden assay interferences (Hoops *et al.*, 2006; Olivier *et al.*, 2005; Pinto *et al.*, 2021).

2.4 Closing statement

The CFR is a unique biodiverse region in South Africa with a range of different environmental conditions. Whilst the actinobacterial biodiversity has been explored to some extent, not many studies have exploited this biodiversity in search of novel MCOs, particularly from rare actinobacteria. This presents a unique opportunity to mine these environments through culture- and sequence-based techniques to access MCO genes that comply with the requirements of the study.

CHAPTER THREE
ACTINOBACTERIAL METATAXOMONICS OF TWO SOIL ENVIRONMENTS

3.1 PREAMBLE

Microorganisms are key players in biochemical geochemical cycles and the interactions between plants and soil microbes are imperative to maintain plant health and soil ecosystems (Le Roux *et al.*, 2017). Several key indicator species have been reported as the dominant phyla in soil communities, of which actinobacteria are always present (Brink *et al.*, 2020; Postma *et al.,* 2016; Weels *et al.*, 2022, Panktratov *et al.,* 2006). A limited number of studies report the actinobacterial diversity associated with the Cape Floristic Region (CFR) soil. Considering the unique environment conditions, ranging from low-nutrient, phenolic-rich ecosystems to organic-rich environments saturated with lignocellulosic material, 16S rRNA gene-based metabarcoding and the determination of soil physicochemical parameters were performed for two CFR regions (soils associated with Rooibos plants in Clanwilliam, and soils associated with a peatland in Springfield, Agulhas) to provide a snapshot overview of these environments and identify the actinobacterial taxa present.

3.1.1 Variations and study limitations

At the time of inception of the study, only a single environment was targeted (Rooibos, Clanwilliam). During the tenure of the project, a rare actinobacterium was isolated as part of a concurrent study (Springfield, Agulhas). It must be noted that as the sampling methods varied between the two studies, no direct comparison can be assumed between the two environments. As such, any assumptions made will be confined to a rudimentary overview of the “actinobiome” and core microbiome for each environment and discussed separately.

3.2 EXPERIMENTAL PROCEDURE

3.2.1 Sampling and determination of physicochemical parameters

3.2.1.1 Environment 1: Rooibos Farms, Clanwilliam

Soil samples were collected from three Rooibos farms (Vaalkrans – 32˚01'07.09" *S* 18˚54'39.34"; Geelland – 32˚02'08.74" *S* 18˚53'01.48" and Muggiesdraai – 32˚00'28.50" *S* 18˚52'41.50") under the ownership of one farmer in the Clanwillian region, Western Cape, South Africa (Figure 3.1). Sample collection took place in August 2016 at the end of the winter season in South Africa.

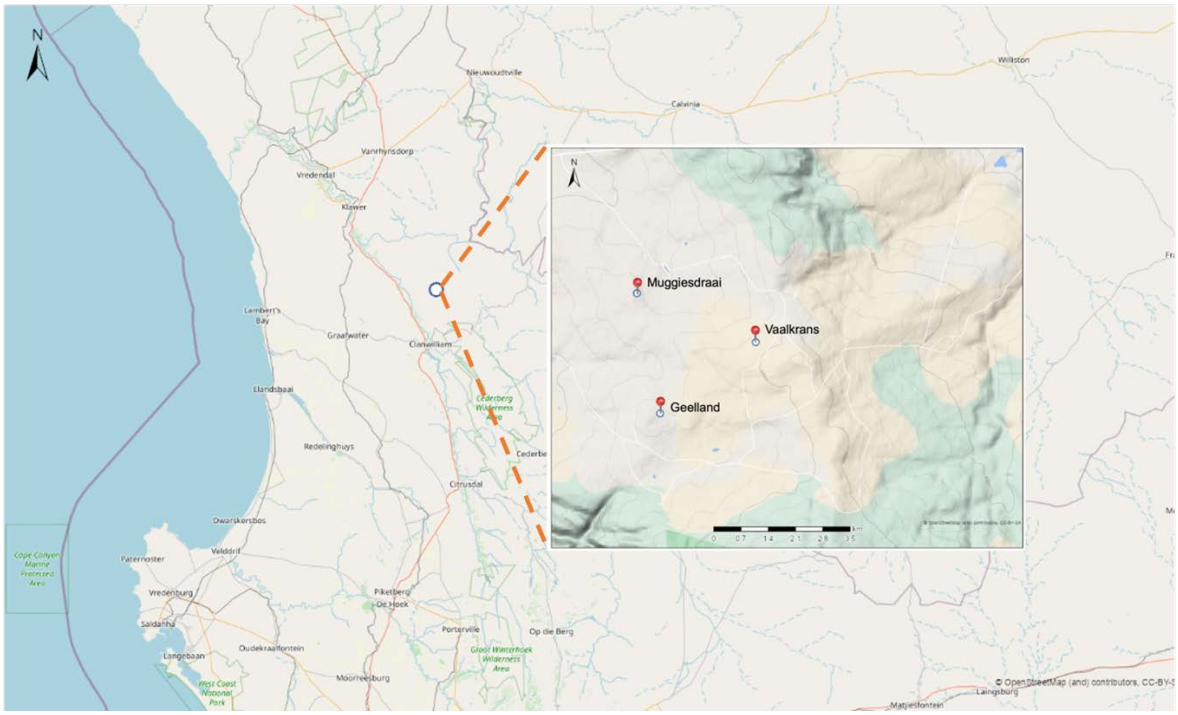


Figure 3.1: Location of the three Rooibos farms ((Vaalkrans – 32˚01'07.09" *S* 18˚54'39.34"; Geelland – 32˚02'08.74" *S* 18˚53'01.48" and Muggiesdraai – 32˚00'28.50" *S* 18˚52'41.50") where sampling commenced in the Clanwilliam region, Western Cape (map created with Cape Farm Mapper 3).

Samples were collected from soil around two wild Rooibos plants within 100m of each other (designated "P1" and "P2") at each of the three sites (designated "S1", "S2" and "S3") at depth of 10 cm (Figure 3.2). Duplicate samples were collected. The soils (~ 50 g) were placed into sterile bags and refrigerated immediately after collection in a portable refrigeration unit (kept at 4 ˚C) for transportation and processed within 24 hours.

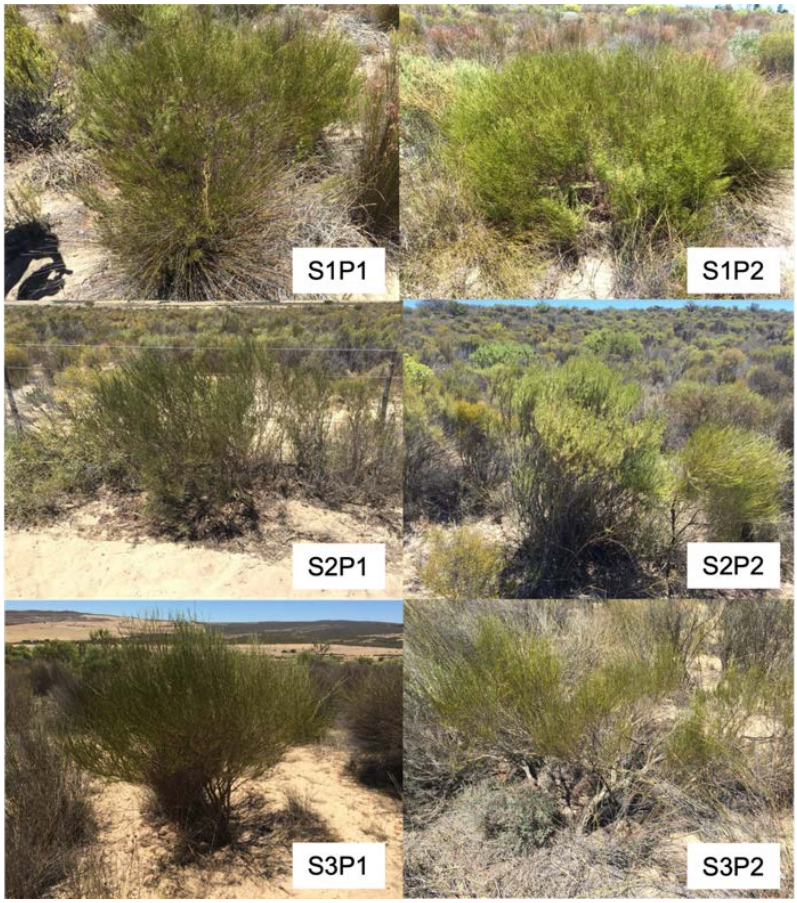


Figure 3.2: Images of the three Rooibos plant selected at Vaalkrans, Geelland and Muggiesdraai for this study. Soil was collected around the base of each plant at a depth of 10cm.

3.2.1.2 Environment 2: Springfield Estate, Agulhas

Soil samples were collected from two sites at a wetland on a privately-owned farm (34˚44'15.3"S 19˚54'38.6") in Springfield, Cape Agulhas (Figure 3.3) in September 2018.

Triplicate samples were collected from two areas within 10m walking distance from each other (designated "Area 1" and "Area 2") at three different depths: top (0 to 5 cm), middle (10-15 cm) and deep (20-25 cm) (Figure 3.4). The soil samples (~ 50 g) were placed into sterile bags and immediately transferred to a portable refrigeration unit (kept at 4 ˚C) for transport and processed within 24 hours.

Figure 3.3: Location of sampling sites at Springfield Estate (34˚44'15.3"S 19˚54'38.6") (map drawn with Cape Farm Mapper 3).

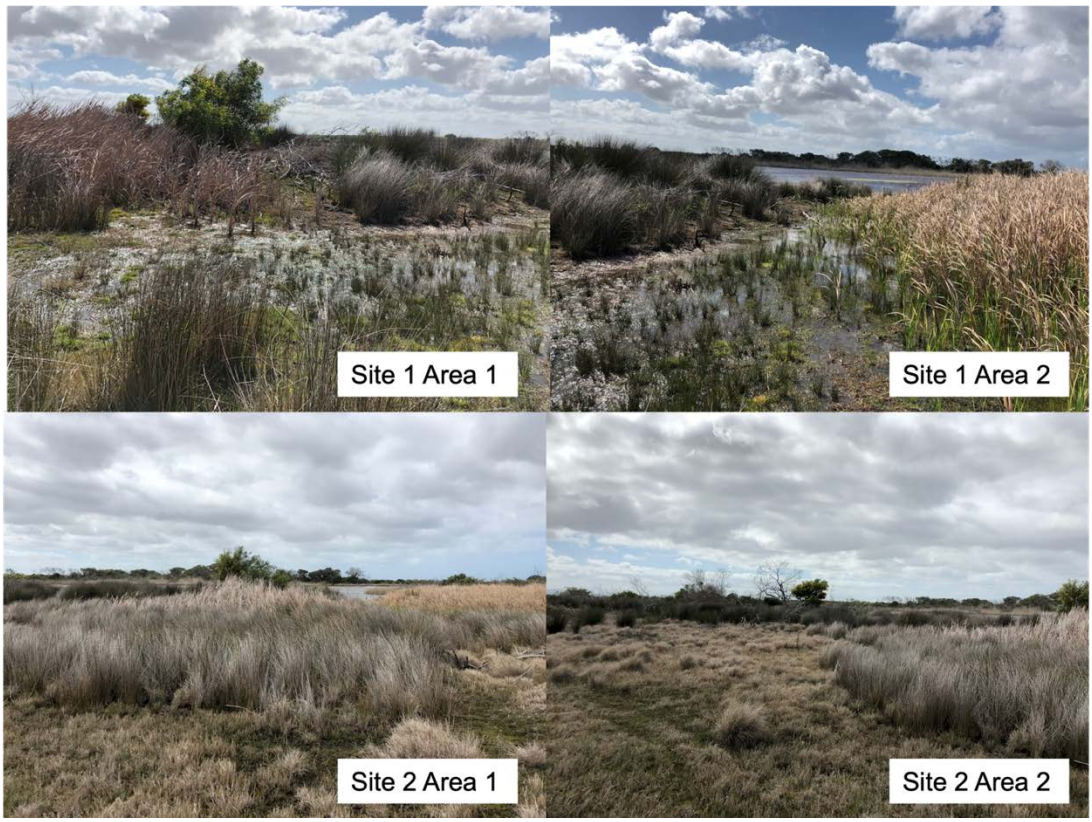


Figure 3.4: Images of the two sites selected at the Springfield Estate for sampling.

3.2.1.3 Physicochemical determination

Bulk samples (~1 kg) for each sampling point were also collected for physicochemical analyses. Samples were submitted to BEMLAB (Strand, South Africa) for a full soil analysis, including soil pH, resistance, phosphorus, potassium, sulphur, carbon, boron, cations, and metals.

3.2.2 Extraction of metagenomic DNA

Metagenomic DNA (mDNA) was extracted using the DNeasy Powersoil DNA Isolation Kit (QIAGEN), according to the manufacturer's instructions with a single amendment: 0.5 g of soil was used, instead of the kit's prescribed 0.25 g to improve DNA recovery.

3.2.3 Amplification and sequencing of actinobacterial-specific 16S rRNA genes

The amplification of the 16S rRNA gene was performed using the method described by Schäfer *et al.* (2010). An actinobacterial-specific 16S rRNA gene primer pair (Com2xf: 5'-AAACTCAAAGGAATTGACGG-3'; Ac1186r: 5'-CTTCCTCCGAGTTGACCC-3') was synthesised by Integrated DNA Technologies (Whitehead Scientific, RSA). A typical PCR reaction consisted of 1x KAPA Taq ReadyMix (containing 1.5 mM $MgCl_2$ and 0.2 mM of each dNTP), 0.2 µM of each primer, 1 µL of template (approximately 10 ng of metagenomic DNA) and water to a final volume of 25 µL.

The amplification program of 25 cycles was started by an initial denaturation step at 95°C for 3 minutes, followed by denaturation at 94°C for 30 seconds, an annealing gradient with temperatures between 51.6°C – 60.2°C and an extension step at 72°C for 30 seconds. A final extension step of 72°C for 15 minutes was performed. Negative PCR controls were included, containing all the components of a typical reaction except for template DNA (water was used instead). Genomic DNA from *Streptomyces polyantibioticus* (SPR^T) was used as a positive control.

Amplicons were analysed by electrophoresis on a 1% (w/v) agarose gel prepared in 1× TAE (containing 10 µg/mL ethidium bromide) and visualised under UV light.

The unprocessed mDNA was submitted to the Molecular Research DNA Laboratory (MrDNA, Shallowater, Texas). MrDNA performed amplicon sequencing using a custom 20 000-read library preparation protocol. PCR for the library prep (bTEFAP® amplicon sequencing with custom barcodes; adaptor sequence: 5'-AAACTCAAAGGAATTGACGG-3'; barcodes: Appendix A) was performed using the actinobacterial-specific 16S rRNA primers (Com2xf: 5'-AAACTCAAAGGAATTGACGG-3'; Ac1186r: 5'-CTTCCTCCGAGTTGACCC-3'). The library was sequenced on the Illumina MiSeq platform (2x300bp chemistry (V3 reagents); paired-end; minimum output of 15 000 – 20 000 reads per library).

3.2.4 Metataxonomics: Data processing and analyses

Pre-processed reads were received from MrDNA. The quality of the pre-processed reads (demultiplexed, barcodes, and indices removed) was determined using FastQC. Read processing and Operational Taxonomic Unit (OTU) picking was performed using QIIME v1.9.1 (Caporaso *et al.,* 2010) installed on a miniconda3 Python environment. Quality filtering (minimum Phred score: 20) and trimming of reads (reads below 50bp) were executed prior to combining reads into a single sequence library. Chimeric sequences were removed using usearch. Open-reference OTU picking was performed using the EZBioCloud 16S rRNA v1.5 database (Yoon *et al.,* 2017) that was formatted for compatibility with QIIME. The resulting OTU file (.biom) was filtered to include only members of the phylum *Actinobacteria* to remove any low-abundance contaminants.

The MicrobiomeAnalyst 2.0 platform (Lu *et al.*, 2023) was used to determine the alpha diversity and core microbiome using the default parameters. The OTU files uploaded were rarefied to the minimum library size to reduce the variability in sample depth and the sparsity of the data between libraries.

3.3 RESULTS AND DISCUSSION

Determining the physicochemical parameters of a given environment is useful to understand the influence different environmental conditions have on the microbial community present. Various factors that influence microbial diversity include soil composition/type (sand, loam or clay), the differing levels of macro- and micronutrients (Na^+, K^+, P, Mn, C and Fe^+, to name a few) and resistance. Bulk soil samples collected at the Rooibos and Springfield sites were submitted to BEMLAB (Strand, Western Cape) to determine the physicochemical parameters of the soil at the time of sampling.

To identify the actinobacterial taxa present at the time of sampling, metabarcoding analyses was performed. Metabarcoding is a useful tool to determine the microbial diversity of a given environment. The actinobacterial constituents of the sampling sites were determined using actinobacterial-specific 16S rRNA gene primers. All samples were screened for the presence of a 270bp amplicon prior to submission of the mDNA for sequencing. OTUs were picked from the pre-processed reads obtained from the sequencing provider (MrDNA) and the major actinobacterial taxa from each environment was determined.

3.3.1 Environment 1: Rooibos

The physicochemical parameters were determined for soils around each Rooibos plant from the three sampling sites (Table 3.1). The soils were comprised entirely of sand for all sites, with an acidic pH ranging from 4.0 – 5.3. Total organic carbon content was similar for all sites, except for the soil around S3P1 which had the lowest carbon percentage of 0.31%. In fact, S3P1 comparatively had lower levels of macro- and micronutrients to the other five sampling points, with the lowest values observed across K, Ca, Mg, Fe, S and P. This agrees with reports of low availability of nutrients typical for this region (Muofhe & Dakora, 2000; Richards *et al.*, 1995).

Table 3.1: Physicochemical parameters determined for soil samples collected from the Rooibos sampling site. Analyses performed by BEMLAB (Strand, Western Cape)

Sample	Soil	pH (KCl)	Resist. (Ohm)	P (mg/kg)	K (mg/kg)	Ex. Cations (cmol(+)/kg)				Cu (mg/kg)	Zn (mg/kg)	Mn (mg/kg)	B (mg/kg)	Fe (mg/kg)	C (%)	Soluble S (mg/kg)
						Na	K	Ca	Mg							
S1P1	Sand	**4.8**	5250	5	26	0.06	0.07	1.10	0.57	0.1	0.1	9.1	0.02	32	0.83	8.10
S1P2	Sand	**4.3**	4710	8	23	0.06	0.06	0.94	0.53	0.1	0.2	5.1	0.11	21	0.77	8.60
S2P1	Sand	**4.4**	1230	6	36	0.14	0.09	0.86	0.65	0.1	0.1	6.9	0.02	53	0.88	11.71
S2P2	Sand	**4.6**	2780	7	50	0.08	0.13	0.85	0.65	0.1	1.4	9.3	0.12	29	0.93	8.09
S3P1	Sand	**5.3**	4680	3	15	0.06	0.04	0.32	0.32	0.0	0.1	3.1	0.22	15	0.31	6.50
S3P2	Sand	**4.0**	3930	4	23	0.06	0.06	0.61	0.39	0.0	0.0	3.1	0.05	43	0.86	9.87

Good sequencing depth was obtained, and the samples were rarefied to the minimum library size (Figure A1: sample sequence size: 62750; Table A2: Good's coverage: 0.99). Since soil composition is generally similar across all the sampling points, it is no surprise that the major actinobacterial constituents (absolute abundance) across all the sites appear to be similar (Figure 3.5). The top five major contributing taxa at family level are *Mycobacteriacea* (14-28%), *Pseudonocardiaceae* (11-29%), *Frankiaceae* (3-10%), *Geodermatophilaceae* (3-21%) and *AF499716_f* (3-26%). Less prominent taxa include members of the families *Microbacteriaceae*, *Micromonosporaceae*, *Nocardiodaceae, Acidimicrobiaceae* and *Nakamurellaceae* (relative abundance; Table A2).

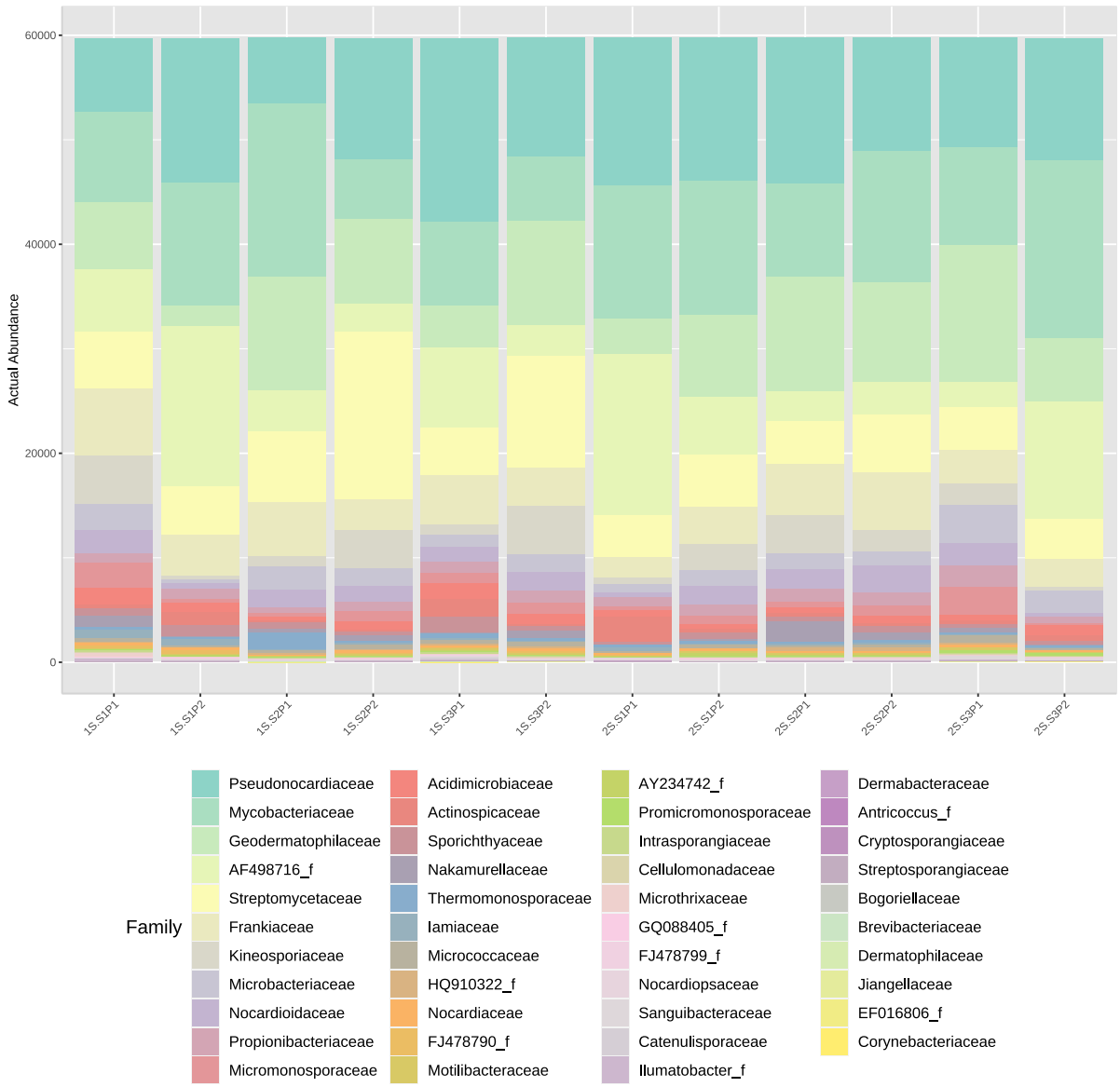


Figure 3.5: Representative taxa (absolute abundance) in soil samples from the Rooibos environments at the family level (rarefied to minimum library size; 1S – replicate 1; 2S – replicate 2).

Overall, low actinobacterial diversity was observed (Shannon Diversity Index < 5) for all soils, regardless of site and plant association (Figure 3.6), and the dominant proportions of the actual abundance are distributed between five families. This is not surprising as several studies have shown that members of of the phylum *Actinobacteria* are among the three dominant phyla used as indicator taxa often detected in soil microbial community studies, suggesting they play an important role in shaping community structure and variation in an environment (Leff *et al.*, 2015; Postma, 2016; Skene, 1998).

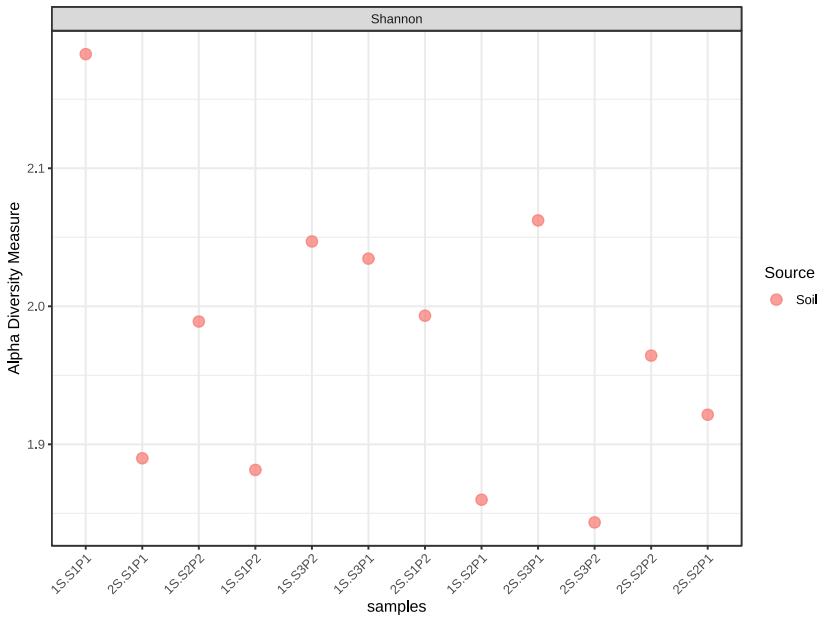


Figure 3.6: Shannon diversity index for the actinobiome associated with Rooibos soil between the three sampling sites (1S – replicate 1; 2S replicate 2).

The findings in this study are also consistent with those reported by Brink *et al.* (2020) where the three dominant phyla were *Acidobacteria, Actinobacteria* and *Proteobacteria* when assessing the bacterial community structure associated with natural and commercially grown Rooibos, and no significant differences in community structure was identified between the two.

3.3.2 Environment 2: Springfield

The physicochemical parameters of the soils collected from the Springfield site were determined for two sites at three depths [0-5cm (top), 5-10cm (middle) and 10-15cm (deep)]. In contrast to the soils associated with the Rooibos plants, the wetland soils consisted of loam at Area 1 and clay at Area 2 (Table 3.2).

The pH ranged from slightly acidic to neutral (6.3-7.1) and a higher level of micro- and macronutrients were available compared to the Rooibos environment, particularly Ca, K, Na, Fe, soluble S and %C. The higher levels of cations (5 to 10-fold higher than the sandy Rooibos soil) is likely due to soil composition since the cation exchange capacity (CEC) is often higher in soil containing clay and organic matter (Botta, 2012).

Also of note are the lower resistance values (10-fold lower compared to the Rooibos environment) measuring between 70-120 ohm. Resistance is an indicator of salinity (soils with resistance < 300 ohm are considered saline). High levels of salinity are anticipated considering the region previously underwent marine transgression and regression events (Hendey, 1983), as well as its proximity to Soutpan (a salt pan, which has a salinity of 16-68 g/kg of salt; Silberbauer & King, 1991).

Good sequence coverage was obtained, and sequences were rarefied to the minimum library size (Figure A2: sample sequence size: 23665; Table A3: Good's coverage: 0.99). Given the soil environment, a larger actinobacterial diversity was expected for the Springfield sites compared to the Rooibos sites, and this is reflected in the abundance profile. Major taxa contributing to the actinobiome are members of the families *Mycobacteriaceae* (5-35%), koll13_f & AKIW874_f (class *Acidimicrobiia;* 4-21% and 2-31%, respectively), *Pseudonocardiaceae* (3-18%) and *Nocardioidaceae* (1-23%). Minor taxa include *Micromonosporaceae, Geodermatophilaceae, Nocardiopsaceae* and *Intrasporangiaceae.*

Table 3.2: Physicochemical parameters determined for soil samples collected from the Springfield sampling site. Analyses performed by BEMLAB (Strand, Western Cape)

Sample	Soil	pH (KCl)	Resist. (Ohm)	P (mg/kg)	K (mg/kg)	Ex. Cations (cmol(+)/kg)				Cu (mg/kg)	Zn (mg/kg)	Mn (mg/kg)	B (mg/kg)	Fe (mg/kg)	C (%)	Soluble S (mg/kg)
						Na	K	Ca	Mg							
A1 Top	Loam	**7.1**	120	12	262	4.6	0.67	2.95	3.04	0.6	1.7	95.8	3.45	509	2.28	54.03
A1 Mid	Loam	**6.3**	130	7	250	4.14	0.64	1.7	2.57	0.3	0.3	55.1	2.46	80	0.63	71.64
A1 Deep	Loam	**6.3**	110	4	248	4.61	0.63	1.86	2.63	0.3	0.3	79.4	3.44	74	0.69	85.7
A2 Top	Clay	**7.1**	110	7	163	4.58	0.42	3.21	3.33	0.4	1.8	25.2	5.08	301	2.31	25.31
A2 Mid	Clay	**6.8**	80	5	248	8.45	0.63	3.14	4.2	0.1	0.2	6.4	2.56	43	1.45	121.08
A2 Deep	Clay	**6.7**	70	5	343	12.17	0.88	3.93	5.79	0.3	0.2	7.5	3.09	27	1.22	177.36

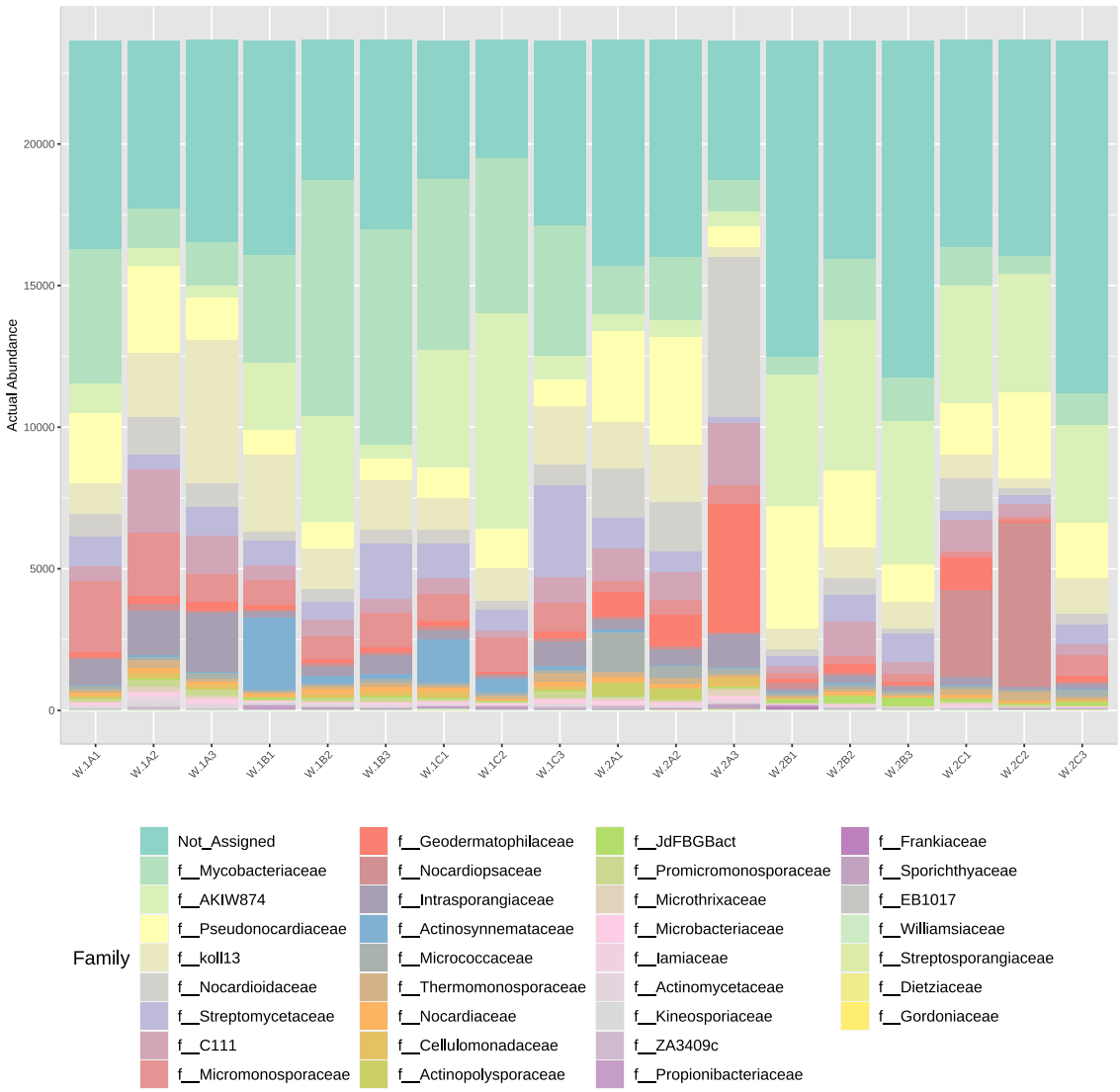


Figure 3.7: Major taxa observed in the Springfield samples for two sites at family level (key: #1*letter*#2 where A – top, B – middle, C – deep; #1 = site 1 or 2; #2 = replicate 1, 2 or 3)

Based on the Shannon Diversity Index, high actinobacterial diversity was observed for most samples across the two sites, with the exception of a few deeper sampling depths. While actinobacterial diversity has been reported to decrease with sampling depth (Aksenov *et al.,* 2021), this decrease may also be due to a number of factors. PCR bias, for example, may lead to unequal amplification of target genes (Acinas *et al.*, 2005). Furthermore, actinobacterial-specific primers have a varying success range of 2 – 87% when amplifying environmental DNA, along with the presence of environmental contaminants which may influence bias by affecting DNA extraction or PCR amplification (Ludermann & Conrad, 2000; Peters *et al.*, 2000, Stach *et al.*, 2001; Stach *et al.,* 2003). Actinobacteria have been found to be present in a range of different oligo-mesotrophic,

ombrotrophic and tundra peatlands (Pankratov *et al.*, 2006). Several of the detected taxa have been reported to play a role in soil. *Geodermatophilaceae*, for example, produce extracellular enzymes that enable their survival in a wide range of soil environments (Montero-Calasanz *et al.*, 2022).

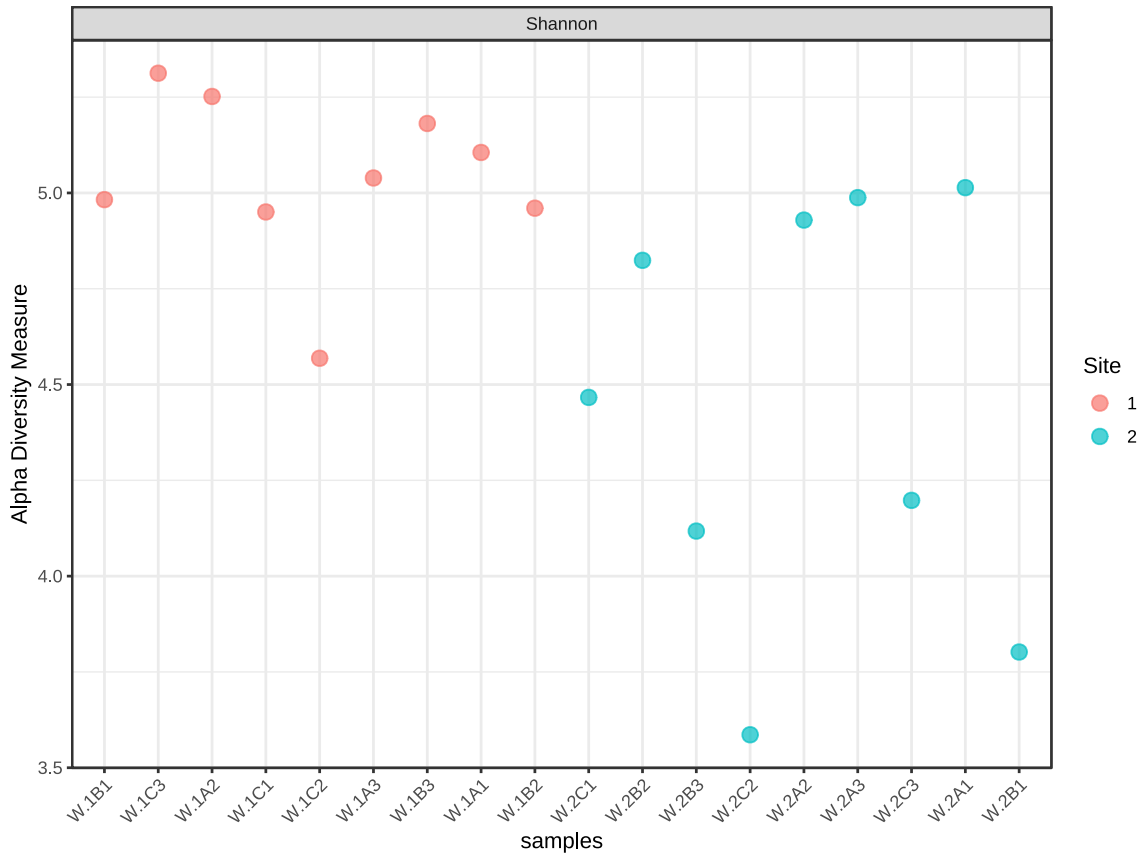


Figure 3.8: Shannon diversity index for actinobiome associated with soil between the three Springfield sampling sites (key: #1*letter*#2 where A – top, B – middle, C – deep; #1 = site 1 or 2; #2 = replicate 1, 2 or 3).

The dominant family *Mycobacteriaceae* has also been extensively reported in peat soils, sphagnum bogs, as well as water run-off from natural and drain peatland (Iivanainen *et al.*, 1999; Iivanainen *et al.*, 1997; Kazda, 2000). They have also been implicated to play a key role in protection against metal toxicity in the immediate environment (Falkinham, 2009). Interestingly, all Springfield samples have rather high Fe^{+} concentrations, and the proliferation of nontuberculous *Mycobacteriaceae* have been shown to be induced by the presence of iron oxide significantly *in vitro* (Glickman *et al.,* 2020) so there is a possible link between the Fe^{+} levels in the soil and the proportion of *Mycobacteriaceae* detected in these samples.

3.4 CLOSING REMARKS

While the approach taken in this study was simple, and only a snapshot of the major taxa represented in the Rooibos and Springfield samples, it is still useful to determine the biotechnological potential of the actinobacterial constituents in these sampling sites, in the context of this study.

The core predicted microbiome (Figure 3.9) for each of these environments, along with the abundance data (Figures 3.5 and 3.7), highlighted several actinobacterial taxa known to produce multicopper oxidases.

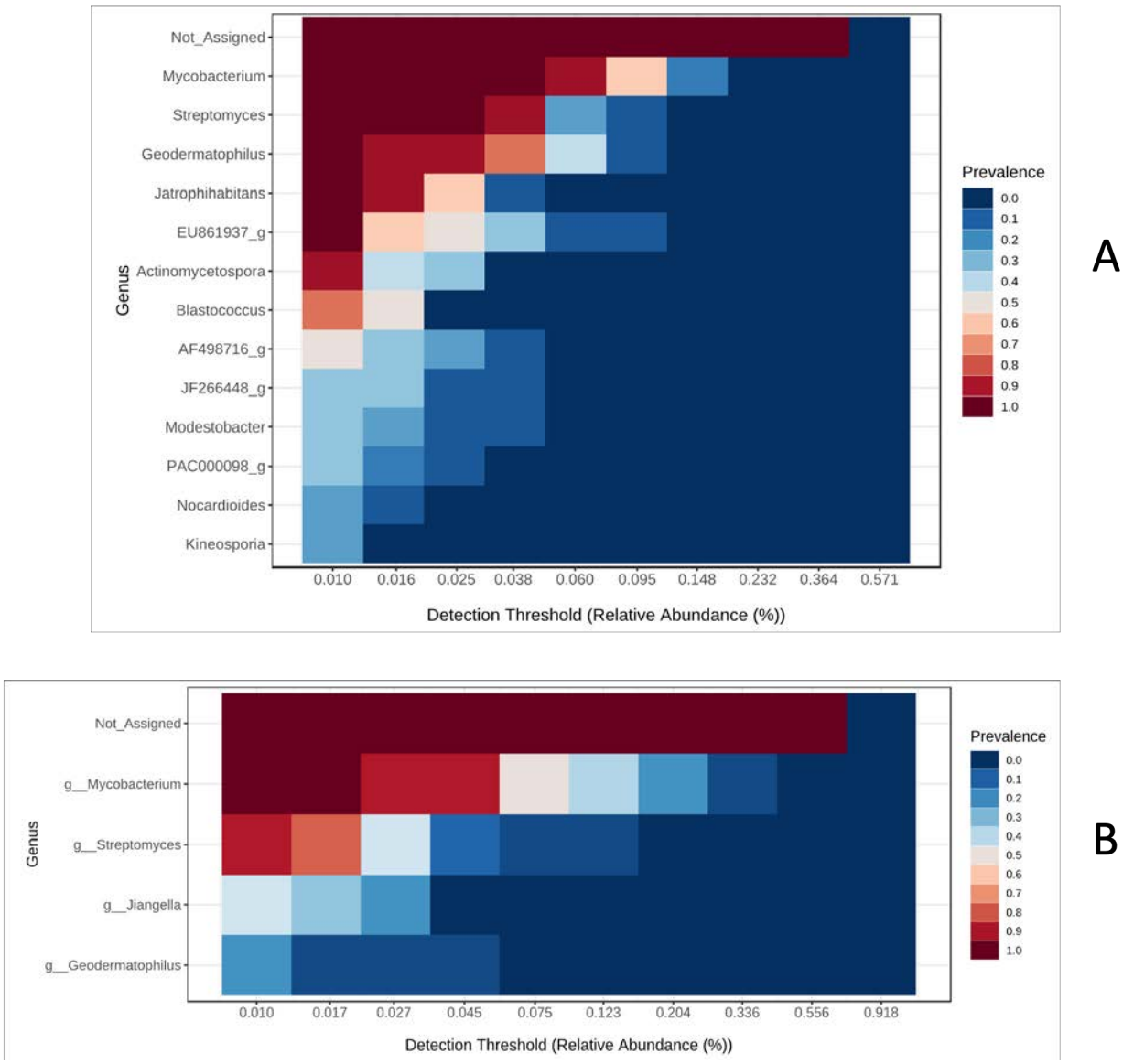


Figure 3.9: Predicted core microbiomes of the Rooibos (A) and Springfield (B) environments.

These findings are not without limitations. While providing a holistic overview of actinobacterial taxa present in these samples, according to a number of studies, the use of actinobacterial-specific primers intrinsically introduces bias due to preferential amplification of specific taxa (Stach *et al.,* 2003; Acinas *et al.,* 2003). Additionally, 16S rRNA gene sequence databases are updated often, and this necessitates the requirement to redesign primers on a regular basis to ensure better coverage (Stach *et al.*, 2003). While not in the scope of this study, in future studies biases can be reduced by performing shotgun sequencing to sequence whole metagenomes and elucidate microbial communities at a greater accuracy than compared to amplicon-based metabarcoding (van der Walt *et al.*, 2017; Greenwald *et al.*, 2017; Vollmers *et al.*, 2017).

When cross-referencing only a select number of major genera represented in the Rooibos and Springfield environments, as well as well-described MCO-producers, *Streptomyces* and *Micromonospora,* with the Laccase and Multicopper Oxidase Engineering Database (LccED) data, 3464 representative MCO sequences were found, across 7 bacterial MCO superfamilies (Table 3.3).

Table 3.3: Examples of MCO sequences in the LccED predicted to be produced by actinobacterial genera also detected in this metabarcoding study (SFAM = Protein Superfamily; 2dMCO: two-domain MCO; 3dMCO: three-domain MCO)

Actinobacterial genus	SFAM 8 (3dMCO)	SFAM 9 (3dMCO)	SFAM 10 (3dMCO)	SFAM 11 (2dMCO)	SFAM 12 (3dMCO)	SFAM 14 (2dMCO)	SFAM 15 (2dMCO)
Mycobacterium	0	73	39	0	665	1	0
Pseudonocardia	0	14	19	0	6	1	3
Frankia	0	16	9	0	6	0	11
Geodermatophilus	0	14	7	0	7	0	0
Nocardioides	0	22	106	0	23	1	0
Jiangella	0	0	1	0	13	0	0
Streptomyces	2	645	385	541	656	6	0
Micromonospora	0	42	31	34	54	1	10

Furthermore, no MCOs produced by actinobacteria isolated from Rooibos have been reported in literature. Together with selective isolation techniques, and the selective pressures these two environments present – low-nutrient, high phenolic in the Rooibos environment, as well as organic-rich lignocellulosic in the Springfield environment –

these sites present a unique opportunity to search for novel MCOs from rare actinobacteria.

CHAPTER FOUR
THE ISOLATION AND IDENTIFICATION OF ACTINOBACTERIA, AND SCREENING FOR MULTICOPPER OXIDASE ACTIVITY

4.1 PREAMBLE

Actinobacteria are Gram-positive bacteria that are widely distributed in terrestrial and aquatic environments (Macagnan *et al.*, 2006). Actinobacteria are able to survive under a variety of environmental pressures such as temperature extremes, and different chemical conditions (which includes alkalinity, high salinity and nutrient-poor environments) (Goodfellow *et al.*, 2018; Shivlata & Satyanarana, 2015). These microorganisms are prolific producers of secondary metabolites, such as bioactive compounds and oxidative enzymes, particularly multicopper oxidases (MCOs), the latter of which is of increased interest due to their application in industrial processes including the degradation of lignocellulosic material and the breakdown of environmental pollutants (Alcalde, 2007; Rodríguez-Couto & Toca Herrera, 2006).

Actinobacteria that are considered "rare" (non-*Streptomyces* species) are not easily cultivated, and limited research on rare actinobacteria producing MCOs have been reported (Seong *et al.*, 2001; Suriya *et al.*, 2016). Since these rare actinobacteria are often underrepresented, they can be considered a good source for the discovery of potentially novel MCOs.

In this chapter, the isolation of rare actinobacteria from the Rooibos environment through selective isolation is described. The isolates were screened for the production of extracellular MCO activity, 'talented' strains identified by 16S rRNA gene sequencing, and the genomes of these organisms were sequenced. The annotated genomes were then explored for potentially novel MCO genes.

4.1.1 Study adaptation

Actinobacterial strains were isolated in a concurrent study from the Springfield environment (not isolated from the soil samples collected in this study). These studies mainly focused on the ability of the actinobacteria to produce antimicrobial compounds. Preliminary genome analyses revealed that a strain belonging to the genus *Actinomadura*, designated "SF1.4" harboured a novel MCO homologous to bilirubin oxidases. As there was overlapping interest with the present study, the raw genome sequence data and the isolate was made available for inclusion in this study (Figure 4.1). The details of the genome reassembly are described in section 4.2.7.

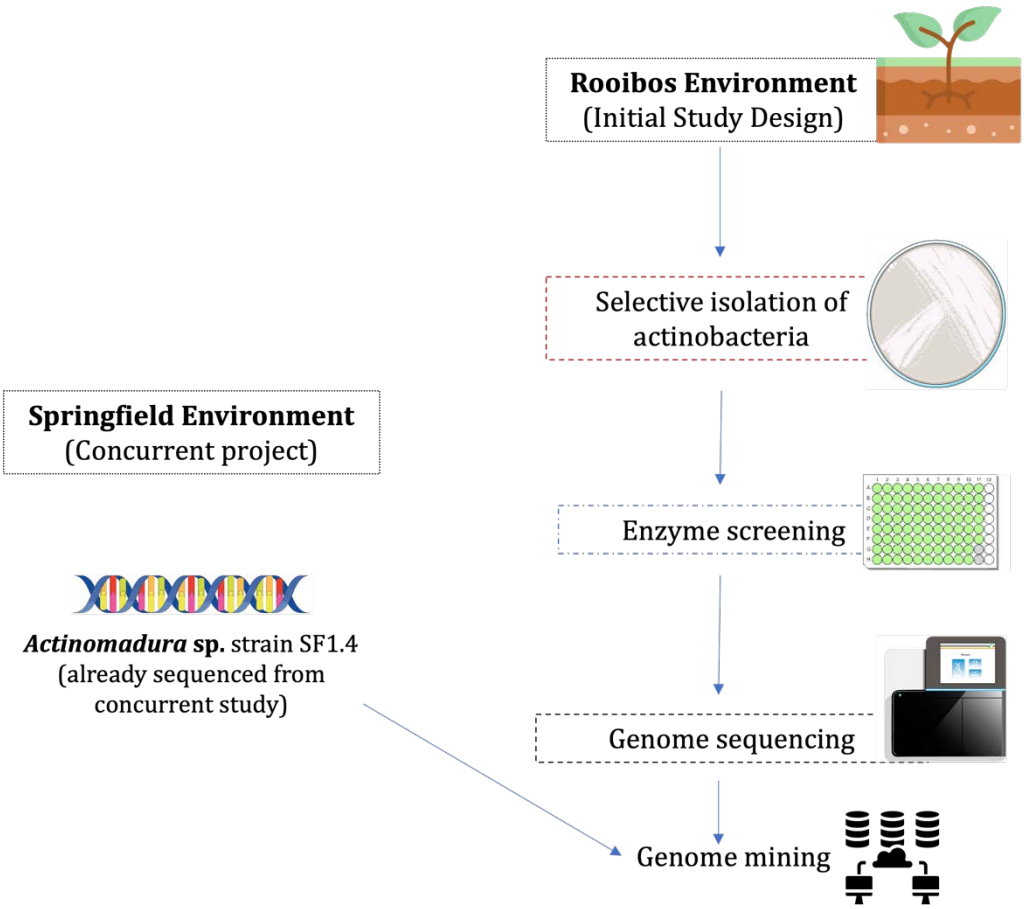


Figure 4.1: Summary outlining project workflow adaptation

4.2 EXPERIMENTAL PROCEDURE

All chemicals were purchased from Merck (South Africa), unless otherwise stated.

4.2.1 Sample pre-treatment

Sample pre-treatment was performed according to strategies described by Hayakawa *et al.* (2008) and Fang *et al.* (2017) for the isolation of rare actinobacteria from Rooibos soil samples. Five grams of soil was heat-treated by microwaving at full power (1000 W) for 15-20 seconds (adapted from Wang *et al.*, 2013). The soils were transferred to a mortar and pestle, an equal amount of $CaCO_3$ (5 g) was added and ground together (Figure 4.2). The pre-treated soil samples were incubated at 30˚C for 10 days.

Figure 4.2: Example of sample pre-treatment by incubating soil with equal amounts of $CaCO_3$

4.2.2 Isolation of rare actinobacteria: Rooibos soils

Actinobacteria were isolated from the pre-treated soils through serial dilution by suspending 1 g of soil in 10 mL of sterile Ringer's solution (10^0 dilution). Subsequent dilutions (1 mL of the previous dilution into 9 mL of sterile Ringer's solution) were made up until a 1×10^{-4} dilution was obtained. One hundred microlitres of each dilution was spread-plated onto various isolation agar media: ISP medium no. 2; R2A, JCM media no.

61, Czapek Solution Agar, Starch-Casein-Nitrate, Humic acid Agar and Rooibos Extract Agar (Appendix B) containing a final concentration of 50 µg/mL penicillin and 50 µg/mL cycloheximide (Glentham Life Sciences, United Kingdom). The isolation plates were incubated at 30˚C for a total of 21 days. Every seven days, the plates were inspected for colonies that displayed atypical *Streptomyces* morphology. The selected colonies were streaked onto media corresponding to their respective isolation media (without antibiotics) to obtain pure colonies. Pure isolates were cultivated in 10 mL liquid cultures (of their respective isolation media) for 5 days at 30˚C, shaking at 160 rpm. Glycerol stocks (20% v/v) of each isolate was prepared and stored at -20˚C until screening for the production of oxidative enzymes. At this stage, the isolates from the concurrent Springfield study were introduced for screening.

4.2.3 Liquid-screening for the production of extracellular MCO activity

The actinobacterial isolates were cultivated by inoculating 50 µL of a glycerol stock into 10 mL Starch-Casein-Nitrate screening media (pH 7.0) and incubated at 30˚C, shaking at 160 rpm. Duplicate flasks were prepared per strain. Samples (500 µL) were taken on days 3, 5, 7 and 10 of cultivation. The supernatant fluid (SNF) was collected by centrifugation at 10 000 rpm for 5 minutes.

Extracellular MCO activity in the SNF was routinely measured using 1 mM ABTS (prepared in 0.05 M sodium acetate, pH 4.0) and 1 mM 2,6-DMP (prepared in 0.05 M Tris-HCl, pH 8.0). Briefly, 50 µL of SNF was transferred to a 96-well microtiter plate. To each well, 150 µL of substrate was added and the increase in absorbance measured over 5 minutes using a Molecular Devices SpectraMax M2 Plate reader (Molecular Devices, San Jose, USA). Laccase from *Trametes versicolor* (TvL, Sigma-Aldrich) was used as a positive control. The negative control included only substrate and buffer.

Enzyme activity (volumetric) was determined using the following formula:

$$\frac{\text{U}}{\text{ml}} = \frac{\frac{\Delta\text{Abs}}{\text{min}}}{\varepsilon} \times \frac{\text{total reaction volume}}{\text{enzyme volume}} \times \text{dilution factor}$$

where:

ΔA/min is the change in absorbance per minute

ε is the extinction coefficient of the substrate in $mM^{-1}cm^{-1}$

dilution factor dilution of sample (if sample was diluted)

Wavelengths for ABTS and 2,6-DMP product formation and their extinction coefficients are listed in Table 4.1.

Table 4.1: Wavelengths (nm) and extinction coefficients for substrates used to detected extracellular MCO activity in this study.

Substrate	λ (nm)	ε (mM^{-1} cm^{-1})	Reference
ABTS	420	36	Eggert *et al.*, 1996
2,6-DMP	477	14.8	Solano *et al.*, 2001

Isolates exhibiting the highest extracellular MCO activity (selection criteria was set to an arbitrary value of 0.05 U/mL of acceptable volumetric activity) were selected for identification using 16S rRNA gene sequence analysis.

4.2.4 Isolation of genomic DNA

The top 20 isolates exhibiting the highest extracellular MCO activity were selected for partial 16S rRNA gene sequence analysis. Isolates were grown in 10 mL ISP2 for 5 days, incubated at 30˚C with shaking at 160 rpm. Cell mass was harvested in 1.5 mL microfuge tubes on day 5 by centrifugation at 10 000 rpm for 5 minutes. Duplicate tubes were prepared for each strain. The culture SNF was completely removed by aspiration with a mechanical pipette.

Genomic DNA was isolated using a modified phenol-chloroform extraction method (Mandel & Marmur, 1968). The cells were resuspended in 400 µL of resuspension buffer

(10 mM Tris-HCl pH 8.0; 1 mM EDTA; 12% w/v sucrose). Dry lysozyme powder (Sigma-Aldrich) was added to each tube to a final concentration of 8 mg/mL and incubated overnight at 37˚C. To each tube, 400 µL of lysing solution (100 mM Tris-HCl pH 8.0; 20 mM EDTA; 300 mM NaCl; 2% w/v SDS; 0.02% v/v β-mercaptoethanol) was added and incubated at 50˚C overnight. Following incubation, 300 µL of phenol:chloroform (2:1 v/v) was added and the sample was mixed on a rotary carousel for 20 minutes at 50 rpm. The tubes were centrifuged at 12 000 rpm for 10 minutes. The aqueous phase was transferred to a clean 1.5 mL microfuge tube and the extraction process was repeated twice with phenol:chloroform (2:1 v/v). After the third extraction, 0.6 volumes of isopropanol were added and mixed gently to precipitate the DNA. The DNA was pelleted by centrifugation at 12 000 rpm for 2 minutes. The SNF was removed and the pellet washed with 500 µL of 75% (v/v) ethanol. The pellet was collected by centrifugation and the wash was repeated. The pellet was dried at 37˚C for 15 minutes. The dry DNA pellet was resuspended by the addition of 400 µL of 1× TE (10 mM Tris-HCl pH 8.0; 1 mM EDTA) and left to resuspend overnight at 4˚C. The next day, 5 µL of RNAse solution (1 mg/mL in 1× TE) was added to each tube and incubated at 37˚C for 1 hour. The RNA was removed by extraction with 100 µL chloroform:isoamyl alchohol (24:1 v/v) and the aqueous phase transferred to a clean 1.5 mL microfuge tube. To each tube, 0.1× volumes of 3 M sodium acetate was added and mixed gently. The mixture was overlayed with 2× volumes of 95% (v/v) ethanol, and the DNA collected by centrifugation. The DNA pellets were resuspended in 500 µL of 1× TE overnight and kept at 4˚C until use.

4.2.5 Amplification of 16S rRNA genes

The partial 16S rRNA gene for the top 20 strains was amplified using universal 16S rRNA gene primers (16S-F1: 5'- AGAGTTTGATCITGGCTCAG-3'; 16S-R5: '5-ACGGITACCTTGTTACGACTT-3'; Cook & Meyers, 2003). A typical PCR reaction consisted of 1× KAPA Taq ReadyMix (containing 1.5 mM $MgCl_2$ and 0.2 mM of each dNTP), 0.4 µM of each primer, 1 µL of template (approximately 50 ng of genomic DNA) and sterile water to a final volume of 50 µL.

The amplification program of 30 cycles was started by an initial denaturation step at 96°C for 2 minutes, followed by denaturation at 96°C for 30 seconds, an annealing temperature

of 56°C and an extension step at 72°C for 30 seconds. A final extension step of 72°C for 5 minutes was performed. Negative PCR controls were included, containing all the components of a typical reaction except for template DNA (water was used instead). Genomic DNA from *Streptomyces polyantibioticus* (SPR^T) was used as a positive control.

The PCR amplicons were analysed on a 1% (w/v) agarose gel (prepared in 1× TAE) containing ethidium bromide (10 µg/mL) and the target amplicon (~ 1.5 kb) was visualised using a Gel Doc XR+ (BIO-RAD, California, USA). The positive amplicons were purified using a MSB Spin PCRapace PCR Purification Kit (Invitek, Berlin, Germany). The concentration of the purified amplicons was measured (minimum concentration: 30 ng/µL) using a Genova Nano Micro-Volume Spectrophotometer (Jenway, Staffordshire, UK).

Amplicons were submitted for Sanger sequencing to Inqaba Biotechnical Industries (Pretoria, South Africa). The raw .ab1 files received were analysed and edited using FinchTV version 1.4.0 (GeoSpiza, Seattle, USA) before submission to EzBioCloud to identify the isolates to genus level (Yoon *et al.*, 2017).

4.2.6 Genome sequencing and analyses

Note: The genome sequence for *Actinomadura* sp. strain SF1.4 was sequenced in a previous study using the Illumina MiSeq platform (2× 300bp, V3 sequencing cartridge). The raw paired-end data (.fastq.gz) was made available for inclusion in this study and reassembled to follow a similar processing pipeline.

4.2.6.1 Genome sequencing: Rooibos isolates

Genomic DNA for the selected Rooibos isolates were submitted to the Central Analytical Facility (CAF, Stellenbosch University, South Africa) for whole genome sequencing. Library preparation and QC was performed, after which the genomes were sequenced on the Ion Torrent S5 platform (on a 530v1 chip, 550bp chemistry, average read coverage > 100×) (ThermoScientific, Massachusetts, USA).

4.2.6.2 Genome assembly

The raw .BAM files were received, pre-trimmed of adaptors and barcodes. The raw sequence quality was determined using FASTQC (version 0.11.9; available online http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Genome assembly was performed using the St. Petersburg Genome Assembler (SPAdes) version 3.15.4 (Prjibelski *et al.*, 2020) on a CentOS high-performance computing cluster (Centre for High Performance Computing, CSIR, Rosebank, South Africa). A PBS-Pro scheduler was used to submit jobs to the processing queue. Each genome assembly used 24 central processing unit (CPU) cores and 128GB of random-access memory (RAM). Error correction was applied prior to assembly. The scripts were also flagged with "—iontorrent" to ensure the correct k-mer sizes were assigned.

The reassembly of the *Actinomadura* sp. SF1.4 Illumina data was performed in the same manner, omitting the "—iontorrent" flag.

4.2.6.3 Genome Quality Check and Annotation

The resultant FASTA outputs were submitted to the Rapid Annotation using Subsystems Technology online tool (RAST; Aziz *et al.*, 2008). A comprehensive genome analysis was also performed on BV-BRC (Gillespie *et al.,* 2011). Post-assembly quality and completeness checks were performed using the CheckM plugin on KBase (version 1.0.18; Parks *et al.*, 2015; Arkin *et al.,* 2018).

4.2.6.4 Protein sequence analysis

Sequences annotated as MCOs or multicopper polyphenol oxidase (MPO) were further analysed using the BLAST function of the LccED (https://lcced.biocatnet.de/; Gräff *et al.,* 2020; Sirim *et al.,* 2011) and UniProt (https://uniprot.org; The UniProt Consortium, 2023).

The presence of signal peptides was predicted using SignalP version 6 (Teufel *et al.,* 2022). Protein domain architecture and putative catalytic sites were determined using the InterPro web server (Payson-Lafosse *et al.,* 2023).

Three-dimensional (3D) models were predicted for the proteins with putative catalytic sites using ColabFold, a Google Colab notebook based on AlphaFold2 (Mirdita *et al.,* 2022; Jumper *et al.,* 2021). Models with a confidence measure below 70% (pLDDT < 0.7) were discarded. The predicted models were visualised in an open-source version of PyMOL (version 2.5.0, Schrödinger, LLC.). PDBsum was used to predict the closest-related structure, and this was used for a rudimentary comparison to the new predicted models (Laskowski *et al.*, 1997).

4.2.6.5 Cloning of MCOs and an MPO

Cloning primers were designed for the two MCOs from strain SF1.4 (designed SF1.4_MCO1 and SF1.4_MCO2) and the MPO from strain 2-8 (designated 2-8_MPO) using SnapGene (version 5.1.; www.snapgene.com). *Nde*I and *Xho*I restriction sites were included for SF1.4_MCO1, *Nde*I and *Hind*III restriction sites for 2-8_MPO, and *Sgf*I and *Pme*I restriction sites were included for SF1.4_MCO2 (Appendix A). Cloning primers were synthesised by Inqaba Biotechnical Industries (Pretoria, South Africa).

The genes of interest were amplified by PCR as follows: a typical PCR reaction consisted of 1× KAPA Taq ReadyMix (containing 1.5 mM $MgCl_2$ and 0.2 mM of each dNTP), 0.4 µM of each primer, 1 µL of template (approximately 10 ng of genomic DNA) and water to a final volume of 25 µL. The amplification program of 30 cycles was started by an initial denaturation step at 96°C for 3 minutes, followed by denaturation at 96°C for 30 seconds, an annealing gradient with temperatures between 51.6°C – 60.2°C and an extension step at 72°C for 30 seconds. A final extension step of 72°C for 15 minutes was performed. Negative PCR controls were included, containing all the components of a typical reaction except for template DNA (water was used instead). The PCR amplicons were analysed on a 1% (w/v) agarose gel (prepared in 1× TAE) containing ethidium bromide (10 µg/mL) and the target amplicons (~ 1.5 kb for SF1.4_MCO1; ~1.6 kb for SF1.4_MCO2; and ~ 0.7 kb for 2-8_MPO) was visualised using a Gel Doc XR+ (BIO-RAD, California, USA). Positive

amplicons were excised from the gel and purified using a NucleoSpin Gel and PCR Clean-Up Kit (Macherey-Nägel, Düren, Germany).

Double-digests were performed on the SF1.4_MCO1 and 2-8_MPO amplicons and the plasmid, pET-20b(+) (Novagen, Merck). The 50 µL reaction consisted of 1 µg of DNA, 1× CutSmart buffer (NEB), 10 units each of *Nde*I and *Xho*I (SF1.4_MCO1)/*Hind*III (2-8_MPO), and sterile water up to a final volume of 50 µL. The same procedure was followed for SF1.4_MCO2 with the following amendments: Flexi®Blend *Sgf*I and *Pme*I restriction enzymes and the pFN18A plasmid (Promega, Wisconsin, USA). The digests were incubated for 10 minutes at 37˚C. Digests were purified using the MSB Spin PCRapace PCR Purification Kit (Invitek, Berlin, Germany).

For the ligation, a vector-to-insert molar ratio of 1:1 was used. The amount of insert was calculated using the following equation:

$$\frac{\text{ng vector}}{\text{kb size of vector}} \times \frac{\text{kb size of insert}}{} \times \text{molar ratio of } \frac{\text{insert}}{\text{vector}} = \text{ng of insert}$$

The ligation reactions consisted of 100 ng of vector DNA, the required amount of insert DNA (36 ng – SF1.4_MCO2; 27 ng – SF1.4_MCO1; 12 ng – 2-8_MPO), 1× LigaFast buffer, 0.1 T4 DNA ligase (Promega) and water up to a final volume of 10 µL. The ligations were incubated at 4˚C overnight.

Ligation reactions were transformed into *Escherichia coli* JM109 chemically competent cells (Promega). Briefly, 5 µL of ligation reaction was added to 50 µL of cells. The cells were incubated on ice for 30 minutes. The transformation reactions were heat-shocked at 42˚C for 20 seconds, followed by incubation on ice for 2 minutes. To each transformation reaction, 450 µL of SOC medium was added and the reaction was incubated at 37˚C for 1 hour, shaking at 180 rpm. After 1 hour, the cells were pelleted by centrifugation (8000 rpm for 2 minutes), 300 µL of the supernatant was removed and the cells were resuspended in the remaining liquid. One hundred microliters were spread onto Luria Agar (LA) plates (containing 100 µg/mL ampicillin) and incubated overnight at 37˚C.

Clones were randomly picked, resuspended in 50 µL of water and colony PCR was performed using the cloning primers to confirm the presence of the insert. Positive clones were inoculated into Luria Broth (LB) (containing 100 µg/mL ampicillin) and incubated at 37˚C overnight, shaking at 160 rpm. Plasmid DNA was isolated from the cells using a GeneJet Plasmid MiniPrep Kit (ThermoScientific, Massaschusets, USA). The plasmids were transformed into *E. coli* BL21(DE3) cells (NEB) and *E. coli* ArcticExpress RP cells (Agilent) following the same transformation procedure as before. Clones were picked and inoculated into LB (containing 100 µg/mL ampicillin) and incubated overnight at 37˚C, shaking at 160 rpm. Glycerol stocks (20% v/v) of the cells were made and kept at -80˚C until further use.

4.2.6.6 Expression of the MCOs and MPO

Starter cultures were prepared by inoculating 50 µL of glycerol stock into 5 mL each of 2YT or ZY media (containing 100 µg/mL ampicillin). The starter cultures were incubated at 37˚C overnight. The following day, 2mL of the starter culture was transferred to a main culture of 25 mL of 2YT or AI media (containing 100 µg/mL ampicillin) and incubated at 37˚C, with shaking at 160 rpm, until an OD_{600nm} of 0.6 was reached. For the 2YT cultures, overexpression was induced using 0.4 mM isopropyl β-d-1-thiogalactopyranoside (IPTG). For the AI cultures, the cultures were simply transferred to their induction temperatures. For each enzyme construct, three cultures were prepared, each incubated overnight at 15˚C (Arctic Express) and 22˚C (BL21).

To determine if overexpression occurred, 20 µL of the overexpressed cultures was mixed with 5 µL of 5× SDS-PAGE sample buffer (0.31 mM Tris-HCl (pH 6.8); 50 % glycerol (v/v); 0.05% bromophenol blue (w/v); 5 mM DTT). The samples were boiled for 10 minutes and loaded onto a 12.5% resolving SDS-PAGE gel. Samples were electrophoresed for 45 minutes at 180V and stained with Coomassie blue staining solution (0.0025% in 45:45:10 methanol:water:acetic acid, v/v) for 1 hour. The gels were destained with destaining solution (5:45:10 methanol:water:acetic acid, v/v) and visualised using a Gel Doc XR+ (BIO-RAD, California, USA).

A rudimentary test of enzyme functionality was performed by scaling up the main culture to 200 mL. The cells from the induced cultures (including a BL21 culture not harbouring any plasmid) were harvested by centrifugation at 10 000 rpm for 5 minutes. The supernatant was removed, and the cells resuspended in 0.1 volumes of cold 10 mM potassium phosphate (pH 7.5). The resuspended cells were sonicated for 10 minutes (30s bursts, 30 second rests). The soluble fraction was separated from the insoluble fraction by centrifugation at 13 500 × *g* for 40 minutes. The soluble fraction was transferred to a clean 50 mL tube, and the insoluble fraction was dissolved in 8 M Urea (in Tris-HCl, pH 8.0).

MCO/MPO activity for the crude enzyme preparations was determined at 340nm using 1 mM ABTS across a pH range of 2.0 – 4.0, and a temperature range of 20-45˚C.

4.3 RESULTS AND DISCUSSION

Actinobacteria are well-characterised for the ability to produce industrially relevant secondary metabolites, such as antibiotics and enzymes. The genus *Streptomyces* is perhaps the most renowned of these – it has been extensively studied because of their ability to produce a wide variety of bioactive compounds and ease of isolation from a range of environments when using standard isolation techniques (Lee *et al.*, 2018). To date, more than 800 valid species has been published (List of Prokaryotic Names with Standing in Nomenclature, LPSN; Parte *et al.,* 2020).

While the possibility of discovering novel secondary metabolites produced by *Streptomyces* exists, considering the ability of actinobacteria to thrive in diverse environmental conditions, the often-unexplored rare actinobacterial genera provide a potentially untapped source of undiscovered secondary metabolites.

4.3.1 The isolation of actinobacteria from Rooibos

The selective isolation of actinobacteria from Rooibos soil was performed using seven different isolation media that were supplemented with penicillin and cycloheximide to inhibit the growth of non-actinobacteria and fungi. The choice of selective media,

including a medium containing Rooibos extract, along with pre-treatments was guided by methods outlined Hayakawa *et al.* (2008) for the isolation of rare actinobacteria.

Selection of isolates was primarily based on morphology. Colonies that were atypical of *Streptomyces* (tough-leathery substrate, with fluffy/powdery spores; Bennett *et al.*, 2018; Lerat *et al.*, 2012) were selected and subcultured onto their respective isolation media. While it is not a cut-and-dried method of identification, considering actinobacterial morphology can differ depending on growth medium composition (Li *et al.,* 2016), this approach resulted in a fair number of isolates that looked morphologically distinct (Figure 4.3).

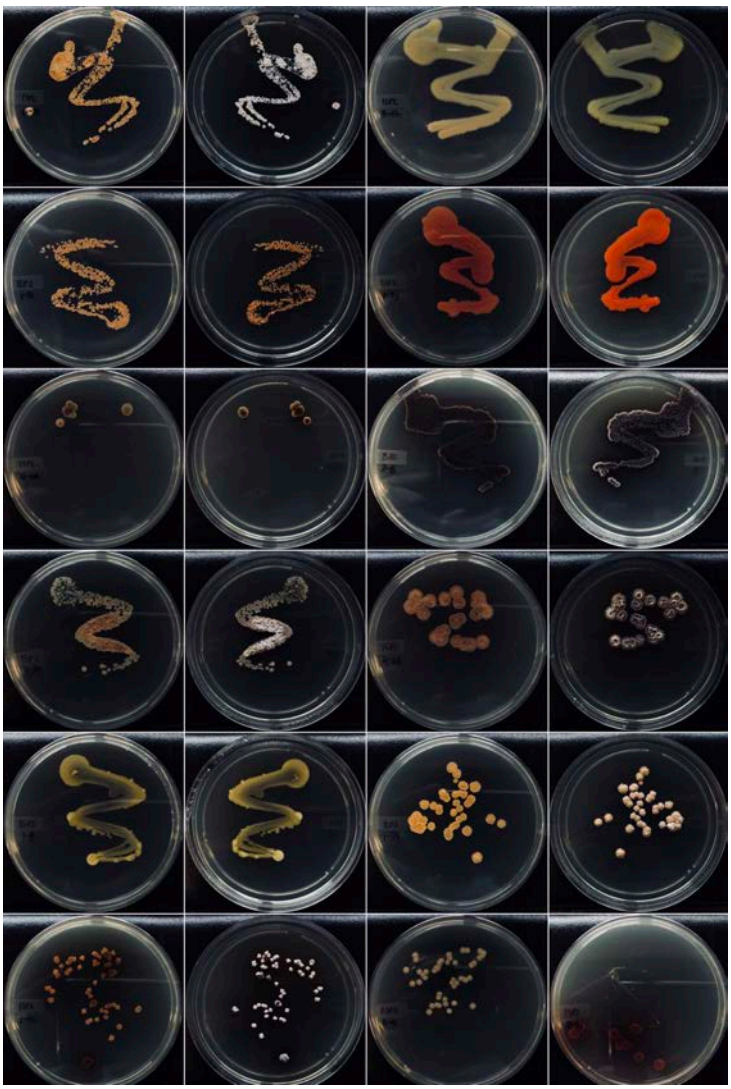


Figure 4.3: Actinobacterial strains isolated from Rooibos grown on ISP2 media to demonstrate range of different morphologies.

A total of 362 actinobacterial isolates were obtained across the six Rooibos samples (Table 4.2). Interestingly, no isolates were obtained using the humic acid agar. This is surprising considering humic acids have been shown to select for actinobacteria (Hayakawa & Nonomura, 1987). Upon subculturing, only isolates that grew well were retained for screening, decreasing the total number to 89 (isolation details for the final number of strains can be observed in Appendix A).

Table 4.2: A summary of total number of actinobacterial isolates obtained from each rooibos soil sample after one round of subculturing onto their respective isolation media.

Isolation Media (agar)	S1P1	S1P2	S2P1	S2P2	S3P1	S3P2
ISP2	15	8	19	16	8	5
R2A	20	3	10	10	9	0
CZ	14	9	16	20	6	1
SCN	4	8	10	14	7	5
0.1% Rooibos Extract	9	1	1	11	5	2
JCM no. 61	15	12	21	24	15	9
Total per site	***77***	***41***	***77***	***95***	***50***	***22***

4.3.2 Screening of actinobacterial isolates for extracellular MCO activity

The actinobacterial isolates were screened for extracellular MCO activity using ABTS (1 mM in 0.05 M sodium acetate, pH 4.0) and 2,6-DMP (1 mM in 0.05 M Tris-HCl, pH 8.0). Activity was monitored over a 10-day period (Figure 4.4).

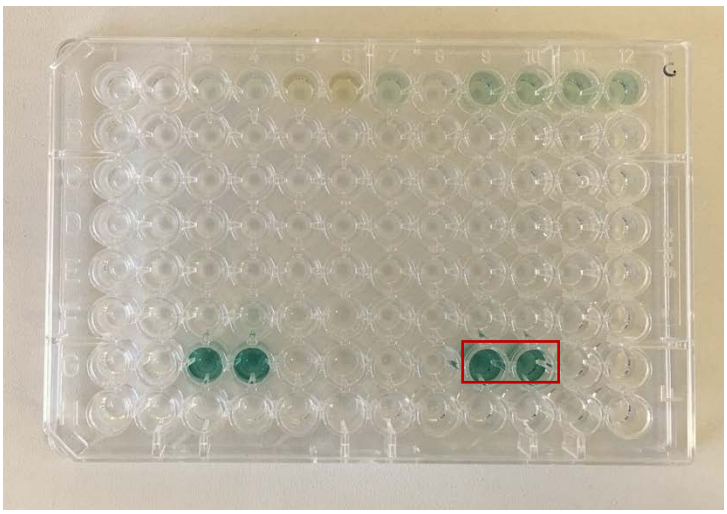


Figure 4.4: An example microtiter plate showing MCO activity from various actinobacterial strains on day 7 of cultivation using ABTS as substrate. Intensity of the green ABTS radical = level of enzyme activity. Enzyme control (*T. versicolor*) is indicated with red block.

Eighty-nine actinobacterial isolates were subjected to screening. Varying levels of volumetric MCO activity was observed (Figure 4.5). The highest activity observed was for isolate 3-33, exhibiting 0.32±0.01 U/mL activity against ABTS on day 3 of cultivation (top producing isolates presented in Appendix A). Nine of the 45 isolates exhibiting activity above 0.05 U/mL were able to oxidise both substrates suggesting that the putative MCO activity occurs at both acidic and alkaline pH, as determined by the substrate used. This is not uncommon – numerous accounts of bacterial MCOs exhibiting broad pH ranges have been reported (Reiss *et al.*, 2013; Lu *et al.,* 2013; Gunne *et al.,* 2012).

For example, the small laccase from *Streptomyces coelicolor* can oxidise 2,6-DMP across a large pH gradient, following a bell-curve distribution with maximal activity at pH 9.4 (Machczynski *et al.*, 2004), whereas fungal MCOs are mostly limited to oxidation of phenolic and non-phenolic substrates under acidic conditions (Alcalde, 2007). This is further corroborated through a comparative study by Reiss *et al.* (2013) where the oxidation of 2,6-DMP and ABTS by five bacterial MCOs, two fungal MCOs and one plant MCO was assessed. All bacterial laccases exhibited activity against ABTS from as low as pH 2.2 to pH 8.1 with > 5% relative activity, and more than 50% relative activity for above pH 3.3. In contrast, the fungal MCOs (*T. versicolor* and *Myceliophthora thermophila*) was largely limited to a pH range of 2.2-5.9, with pH optima for both ABTS and 2,6-DMP below a pH of 3. The plant MCO (*R. vernificera* – the classical laccase), also functioned well across a broad pH range, albeit at a lower optimum for ABTS and 2,6-DMP at pH 3.8 and pH 8.0, respectively. A larger proportion of the putative MCOs from the actinobacterial isolates preferentially oxidised 2,6-DMP. Again, this is suggestive of alkaline-active MCOs – a trait that is highly desirable for application in industrial processes.

For the top 25 potential MCO producers, 16S rRNA gene amplicons were submitted for sequencing. Since there are several isolates exhibiting above 0.1 U/mL activity, selection of the top 25 was based on not only highest level of activity, but also whether the isolate was able to oxidise both 2,6-DMP and ABTS.

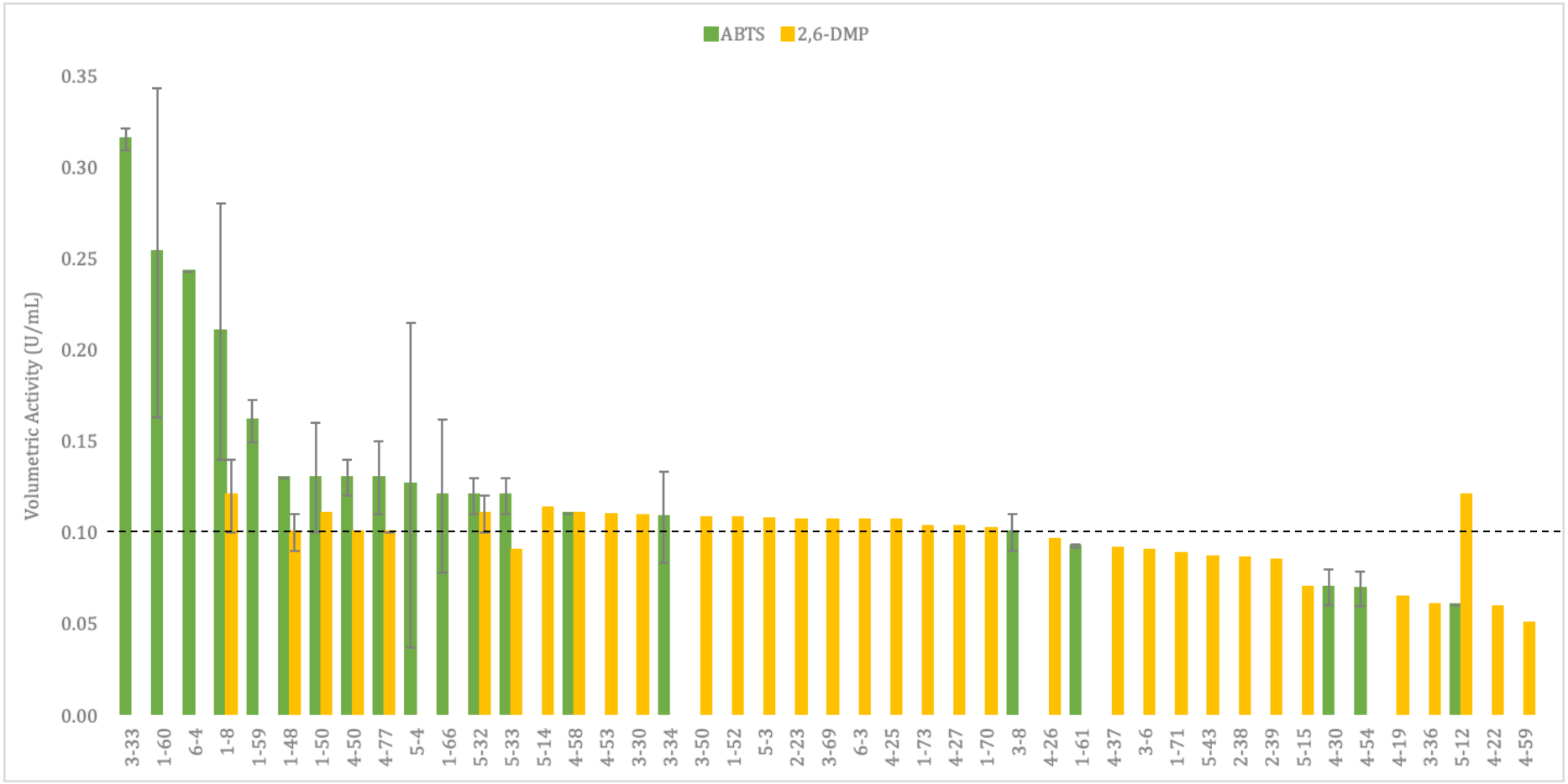


Figure 4.5: Maximum volumetric activity (U/mL) measured over a 10-day period for actinobacterial isolates using 1 mM ABTS and 1 mM 2,6-DMP as substrates. Some isolates exhibited MCO activity at both acidic and alkaline pH (activity against both substrates). The top 25 isolates exhibiting activity above 0.1 U/mL (dashed line) were selected for 16S rRNA identification.

4.3.3 16S rRNA gene sequencing

The 16S rRNA gene of the top 25 MCO producers that exhibited MCO-type activity was amplified and sequenced using Sanger sequencing. Raw sequences were analysed with FinchTV to assess sequence quality and remove and/or edit ambiguous base calls. The edited sequences were assembled and submitted to EzBioCloud for identification to genus level (Yoon *et al.,* 2017).

The majority of isolates were identified as actinobacteria (Table 4.3). Eight of the 25 isolates were identified as rare actinobacteria (*Dactylosporangium, Actinokineospora, Curtobacterium, Modestobacter*, *Microbispora, Leifsonia*, and *Actinomadura*) (Table 4.3).

Table 4.3: The 16S rRNA gene identities of the top 25 isolates from the MCO screening results. Non-actinobacteria are highlighted in red.

Strain	Top-hit taxon	Top-hit strain	Completedness (%)
1-48	*Micrococcus aloeverae*	AE-6^T	95.1
1-50	*Dactylosporangium tropicum*	KB2-4^T	86.4
1-52	*Methylobacterium tardum*	RB677^T	91.0
1-59	*Streptomyces sannanensis*	NBRC 14239^T	95.2
1-60	*Roseomonas rubra*	S5^T	92.3
1-66	LMRQ_s (*Rhodococcus*)	Leaf258	83.2
1-70	*Actinokineospora spheciospongia*	EG49^T	95.5
1-73	*Rhodococcus corynebacterioides*	DSM 20151^T	93.3
1-8	*Curtobacterium flaccumfaciens*	LMG 3645^T	96.3
2-23	*Streptomyces ramulosus*	NRRL B-2714	95.7
3-30	*Streptomyces beijiangensis*	NBRC 100044^T	95.3
3-33	*Modestobacter marinus*	42H12-1^T	94.7
3-34	*Streptomyces scabrisporus*	DSM 41855^T	90.1
3-50	*Methylobacterium tardum*	RB677^T	93.3
3-69	JOFZ_s (*Streptomyces*)	NRRL F-5126	96.1
3-8	JOFZ_s (*Streptomyces*)	NRRL F-5126	89.9
4-25	*Microbispora rosea* subsp. *rosea*	ATCC 12950^T	94.8
4-26	*Streptomyces echinatus*	NBRC 12763^T	57.4 (R5 only)
4-27	*Streptomyces geysiriensis*	NBRC 15413^T	96.1
4-50	*Pseudonocardia xinjiangensis*	AS 4.1538^T	95.6
5-12	*Leifsonia soli*	TG-S248^T	96.1
5-32	*Pseudonocardia bannensis*	YIM 63101^T	90.5
5-33	*Actinomadura nitritigenes*	DSM 44137^T	94.0
5-4	*Roseomonas rubra*	S5^T	91.6
6-4	*Actinomadura nitritigenes*	DSM 44137^T	95.0

To put this into perspective, currently only 11 species of the genus *Modestobacter,* 13 for the genus *Microbispora,* 19 species of *Actinokineospora* and 12 species of *Leifsonia* are validly published (Parte *et al.*, 2020).

Additionally, by cross-referencing these rare genera with the BRENDA database, no records of MCOs or MCO-like enzymes are found. Searching across 7 bacterial superfamilies in the LccED, only 74 records are obtained among 3 400 bacterial MCO sequences represented in 5 superfamilies, further highlighting how underrepresented MCOs are from these rare actinobacteria.

Table 4.4: Summary of bacterial MCOs from rare actinobacteria across 5 superfamilies represented in the LccED.

Superfamily	Group	*Curtobacterium*	*Modestobacter*	*Actinokineospora*	*Leifsonia*	*Microbispora*	*Actinomadura*
9 - Bacterial Bilirubin Oxidase	**3dMCO**	0	0	5	0	3	10
10 - Bacterial CueO	**3dMCO**	0	1	0	10	0	3
11 - SLAC-like (type B 2dMCO)	**3dMCO**	0	0	0	0	7	0
12 - Bacterial MCO	**2dMCO**	1	2	0	8	13	5
14 - Bacterial type B 2dMCO	**2dMCO**	0	0	0	2	1	3

In combination with their putative MCO activity (the level of volumetric activity, as well as activity at both acidic and alkaline pH), underrepresentation in enzyme databases and validly published species - isolates 3-33 (0.32 U/mL), 4-25 (0.11 U/mL) and 5-12 (0.12 U/mL) were selected for further investigation.

4.3.4 Genome sequencing and analyses

Note: At this stage, viability of isolate 4-25 (a *Microbispora* species) was lost, despite concerted efforts to revive it. As such, an additional strain, isolate 2-8 identified as a *Lentzea* species, was selected for inclusion. While 2-8 did not exhibit activity against 2,6-DMP and ABTS, it did show activity against 2,6-dichlorophenol (putative peroxidase activity) and L-DOPA (putative tyrosinase activity) (Appendix A). However, there is known overlapping oxidation activity with tyrosinases, for example, and it is notoriously difficult to assign an MCO a specific name based on activity alone (Reiss *et al.*, 2013). Therefore, the decision was made to include the Rooibos strain, isolate 2-8, from this point forward.

Five genomes were generated with the SPades assembler (version 3.15.4). For each of the draft genomes, a comprehensive genome analysis was performed on BV-BRC (Gillespie *et al.*, 2011).

Draft genome lengths ranged from the smallest in size 3 882 514 bp for the *Leifsonia* sp. isolate 5-12 to the largest for *Actinomadura* sp. isolate SF1.4 at 10 828 139 bp (Table 4.4). The %GC mol content for all genomes were typical of that for *Actinobacteria* (Ventura *et al.,* 2007) ranging from 69.54% to 74.37%. The draft genomes for isolate SF1.4 and isolate 1-70 would generally be considered of poor quality due to the large number of contigs (> 1000). This is a common occurrence when using short-read sequencing that generates short contigs which are often difficult to join (Lischer & Shimizu, 2017). The problem can be overcome by either performing hybrid assemblies in tandem with long read sequencing technologies such as PacBio (Yang *et al.*, 2016), or as more genome sequences become available for a given organism, if a close-related genome is available, reference mapping can be used to assist in filling the gaps (Pop *et al.,* 2004).

Table 4.5: Key characteristics of the draft genomes generated in this study (summarised from the BV-BRC comprehensive analysis and CheckM reports [Appendix A]). No plasmids or partial CDS was detected.

Isolate	SF1.4	2-8	1-70	5-12	3-33
Sequencing Technology	Illumina MiSeq	Ion Torrent S5	Ion Torrent S5	Ion Torrent S5	Ion Torrent S5
Contigs	1 365	238	1 433	105	974
GC content %	72.76	69.54	72.47	68.92	74.37
Genome Length (bp)	10 828 139	9 439 572	7 439 712	3 882 514	4 675 846
Contig N50	33141	98 185	10 384	86 660	9 020
CDS	11 122	9 482	8 508	3 936	5 483
tRNA	80	62	50	44	47
rRNA	10	5	3	3	4
Completeness	99.47	97.65	97.25	99.44	94.35
Contamination	2.85	1.04	1.94	0.0	1.11

To explore potential MCOs that may be present in the genomes of these rare actinobacteria, the protein annotations for each genome was manually searched using the terms “copper”, “putative oxidoreductase” and “laccase”. MCOs were detected in all draft genomes, except for isolate 5-12 (Table 4.6).

Table 4.6: Putative MCOs and MCO-like proteins annotated in the genomes of rare actinobacteria

Annotation	2-8	SF1.4	3-33	1-70	5-12
Multicopper oxidase / Laccase / Putative oxidoreductase	1	2	2	4	-
Multicopper polyphenol oxidase	1	1	-	2	-

An analysis of the predicted protein annotations revealed 13 (in total) potentially novel MCOs and MPOs (MCO-like proteins), an underrepresented enzyme class with only a limited number reported, primarily from metagenome sequencing projects (Chai *et al.*, 2017; Jeon & Park, 2020; Narnoliya *et al.*, 2019; Sharma *et al.*, 2019).

4.3.5 Protein Sequence Analyses

Nine potential MCOs and four potential MPOs were identified through the primary RAST and UniProt annotations. The amino acid sequences were uploaded to InterPro to search for known domain architectures, putative catalytic sites, and predicted activity. Four of the total 13 potential proteins were selected for further analyses based on these criteria. The amino acid sequences for these four proteins are listed in Appendix A.

4.3.5.1 MCOs

For one MCO from isolate 1-70 (designated 1-70_MCO1) and two MCOs from isolate SF1.4 (designated SF1.4_MCO1 and SF1.4_MCO2), cupredoxin domains were detected that were consistent with typical 3dMCOs. These 3dMCOs are usually composed of a single peptide chain of approximately 500 amino acids, with three sequential cupredoxin-like domains (Figure 4.6; Komori & Higuchi, 2015). Copper-binding sites and domain interfaces were also detected. Oxidoreductase activity predicted for all three 3dMCOs, suggesting an intact active site.

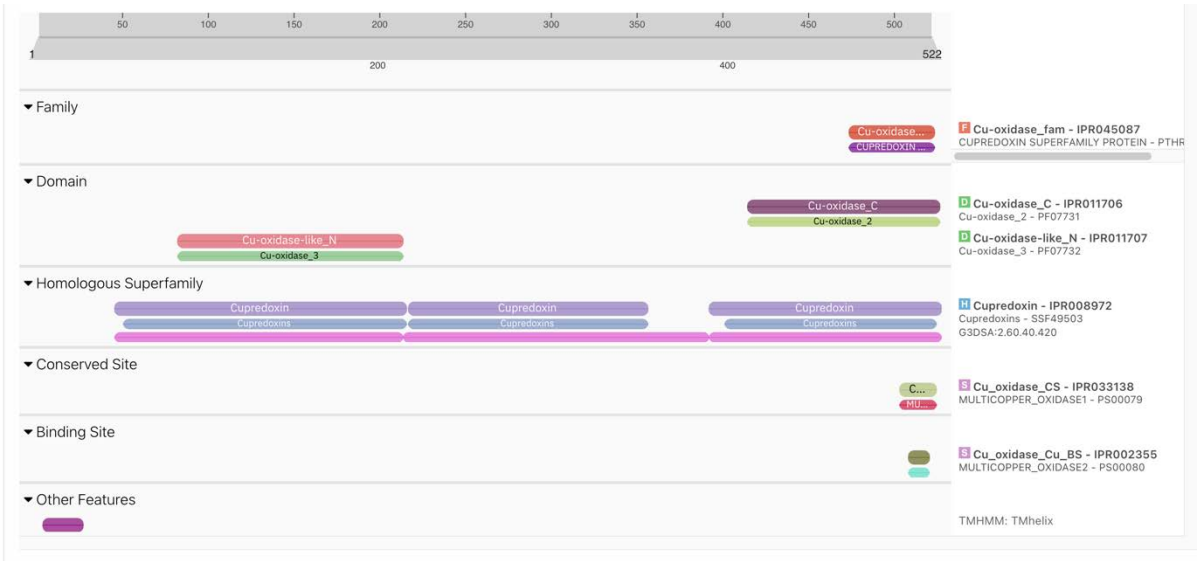


Figure 4.6: InterPro summary of the features detected for 1-70_MCO1. The remaining InterPro summaries are listed in Appendix A.

pSORTb predicts the likelihood of a protein's subcellular localisation. For all three 3dMCOs, a high likelihood of cytoplasmic localisation of 9.97 was predicted (Appendix A).

Twin-arginine translocation (TAT) signal peptides were detected for all three MCOs, between positions 0-25 with a likelihood of 0.9266, 0.8332 and 1 for 1-70_MCO1, SF1.4_MCO1 and SF1.4MCO2, respectively. The TAT pathway, in bacteria, catalyses the export of folded protein from the cytoplasm across the inner membrane (Lee *et al.*, 2009) and the presence of the signal peptide suggests these MCOs are exported from the cell to

the extracellular medium, which could be why activity was observed for these isolates in the screening experiments.

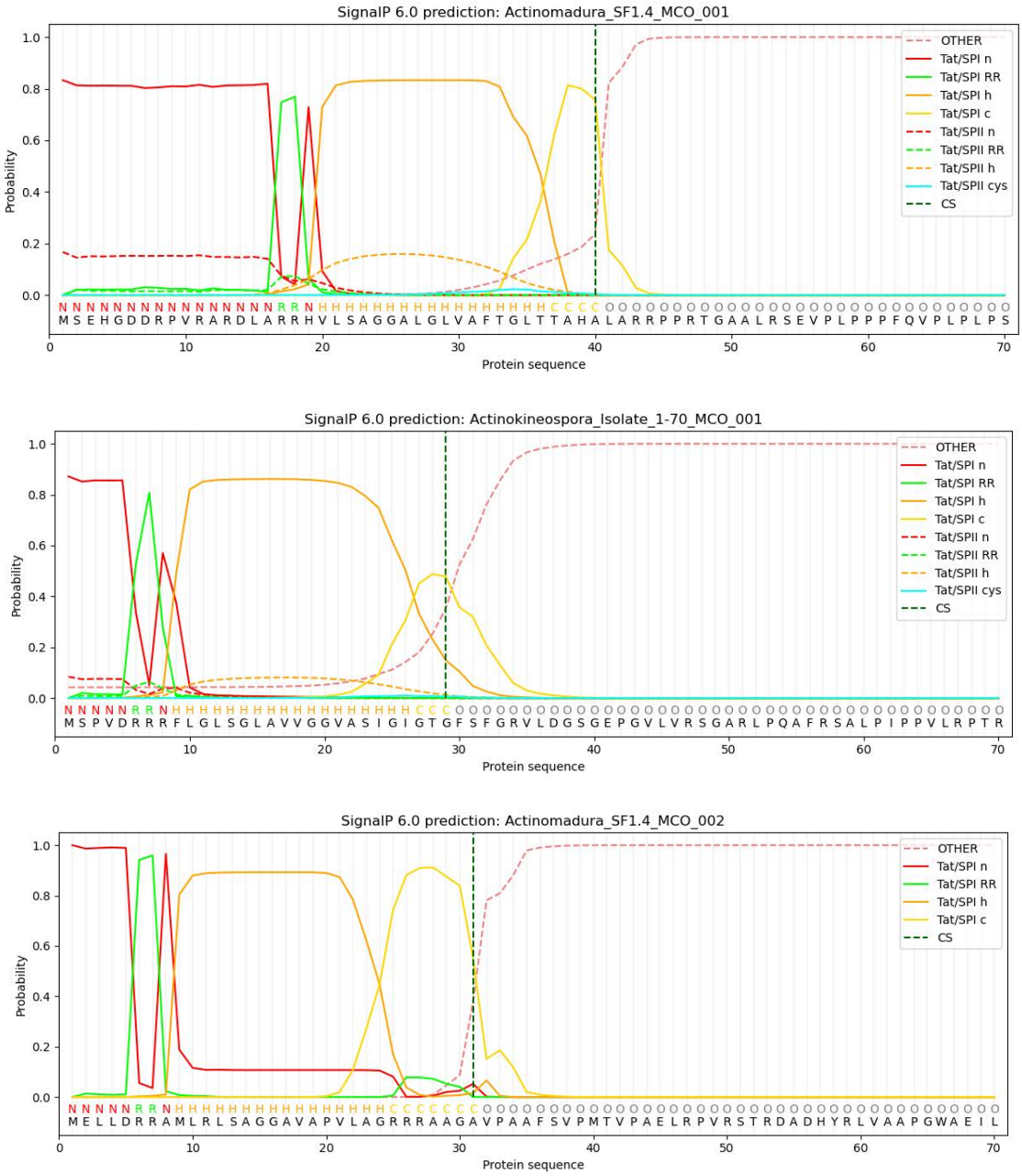


Figure 4.7: Signal peptide predictions using SignalP version 6.0 for 1-70_MCO1, SF1.4_MCO1 and SF1.4_MCO2. Twin-arginine translocation signal peptides were detected for all 3dMCOs selected.

The crystal structure for the *Bacillus pumilis* bilirubin oxidase (PDB ID: 7Z5P, Gihaz *et al.*, 2022) was selected as the reference protein for rudimentary structural visualisation based on having the most overlapping homology for the three predicted models (36-

38%). All predicted models scored above a 90% confidence level. Structural alignment shows that the overall secondary structure was highly similar (Figure 4.8).

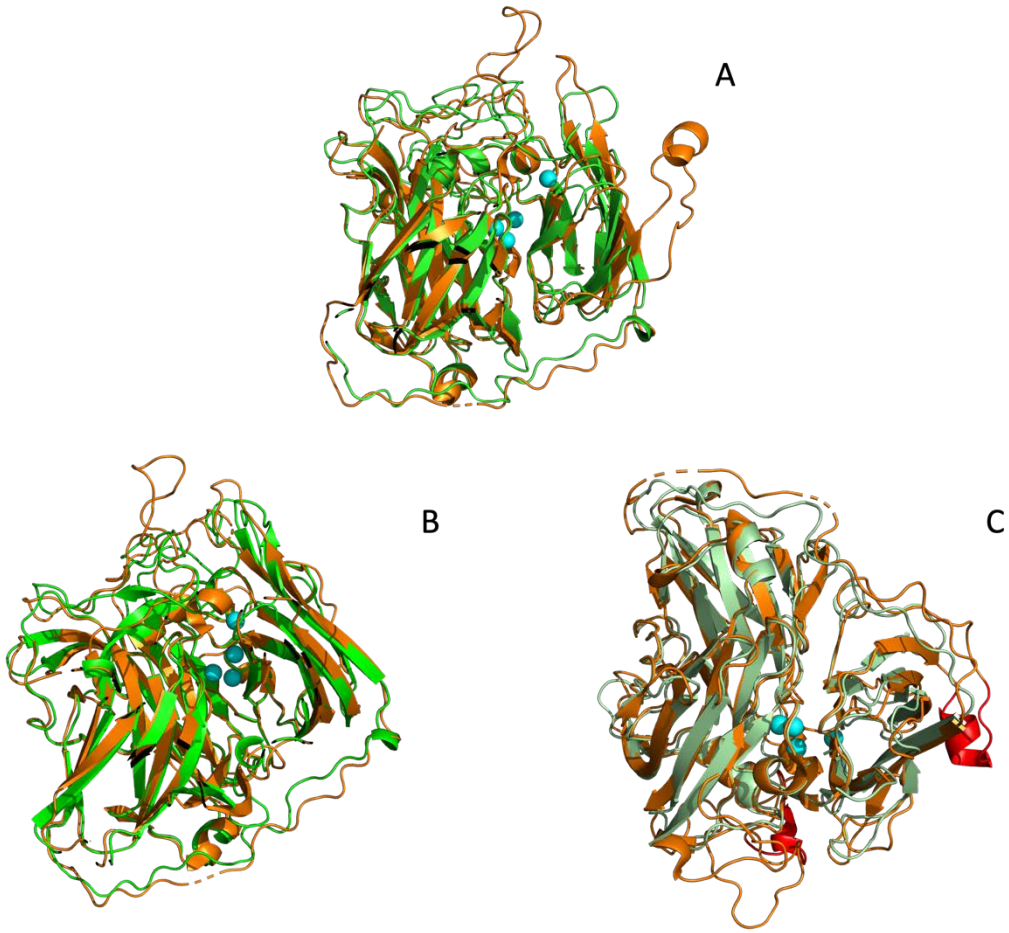


Figure 4.8: Three-dimensional predicted structures (green) of 1-70_MCO (A). SF1.4_MCO1 (B) and SF1.4_MCO2 (C) aligned with the bilirubin oxidase from *Bacillus pumilis* (orange - PDB ID: 7Z5P) showing a high degree in similarity of overall folding. Copper ions are depicted in cyan.

Long loops cover the active site, similar to the *B. pumilis* bilirubin oxidase (Gihaz *et al.*, 2022). Multiple sequence alignment of the signature motifs near the active site showed that the region around the T1 copper and T2/T3 trinuclear cluster (TNC) is conserved (Figure 4.9).

```
Bacillus_pumulis_bilirubin_oxidase_BpBOD      PNEAPPLHEQGYKDTIQAHAGEVIRIIARFV-PYSGRYVWHCHILEHEDYDMMRPMDIIH      510
Actinomadura_SF1_4_MCO_002                    -GGPPLPTERGLKDTVMLPPGGSVRIRVSFDPRYCGDYVYHCHVIDHSSMGMMGRFRVAG      469
Actinokineospora_Isolate_1-70_MCO_001         -NGGPGPYDAGWKDTVDLHPAQAMEVAIRFT-DYPGRFVFHCHNLEHEDMGMMADFTTL-      522
Actinomadura_SF1_4_MCO_001                    -GGPPPPHDAGLKDTVSLRPGEAVEIITRFD-GYRGRYLFHCHNAEHEDMGMMANLEVV-      542
                                               .  *   : * ***:    .  :.:   *   * * :::***  :*.. .**  :
```

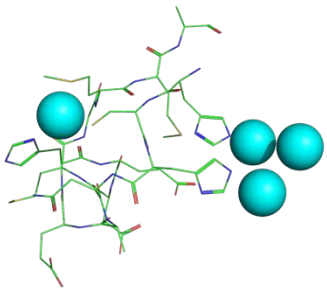


Figure 4.9: Residues coordinated to the active site. The highlighted region in the multiple sequence alignment depicts the conserved M4 signature pattern (H-C-H-X3-H-X3-[AG]-[LM]) (Messerschmidt & Huber, 1990).

4.3.5.2 MPO

For the MPO identified from the genome of the *Lentzea* sp. isolate 2-8 (designated 2-8_MPO), InterPro analyses detected the presence of a YfiH / Cu_oxidase domain homologous to that found in *Escherichia coli*, as well as the first fully characterised MPO, the RL5 laccase (Beloqui *et al.,* 2006). These enzymes typically have amino acid sequences shorter than 300, as is the case with this 2-8_MPO being only 234 aa in length. A putative catalytic site was also detected.

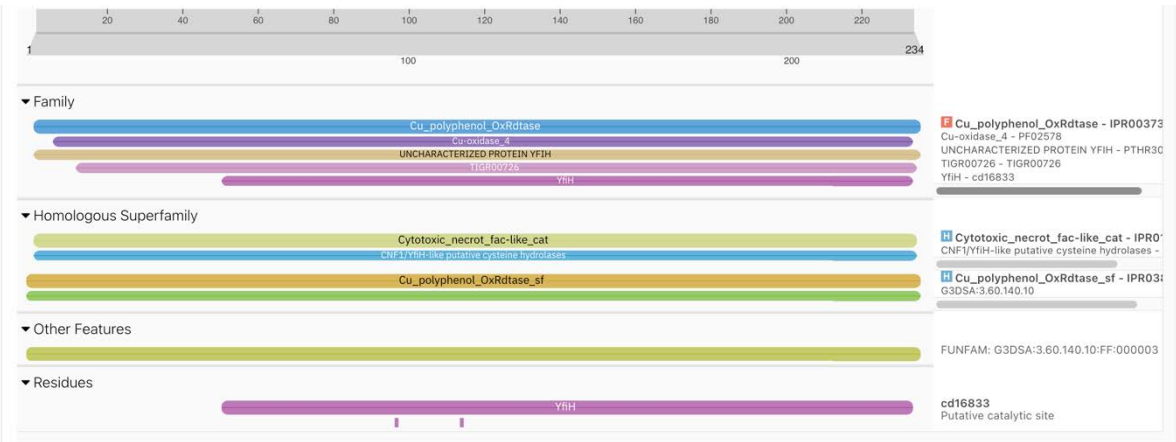


Figure 4.10: InterPro summary of the features detected for 2-8_MPO.

pSORTb analyses predicts, with a likelihood of 7.75 (Appendix A), that 2-8_MPO is a cytoplasmic protein. No signal peptides were detected.

The YfiH crystal structure from *Shigella flexneri* (PDB ID: 1XAF; Kim *et al.*, 2006) representing the protein family Domain of Unknown Function 152 (DUF152) was selected as the reference structure for visual comparison to the predicted protein model of 2-8_MPO. The ColabFold model for 2-8_MPO was scored at a confidence level of 98%. Structural alignment of 2-8_MPO with the YfiH from *S. flexneri* showed a high degree of similarity (Figure 4.11) in overall fold.

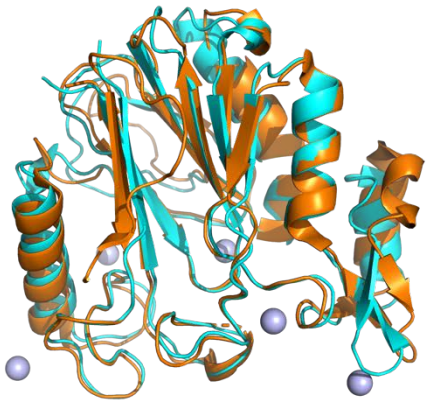


Figure 4.11: Three-dimensional predicted structures (cyan) of 2-8_MPO aligned with the YfiH (orange - PDB ID: IXAF) showing a high degree in similarity of overall folding. Zinc ions are depicted in light purple.

Based on the structural characteristics of the YfiH from *S. flexneri* and the RL5 laccase, it was found that conserved histidine and cysteine residues are maintained at the active site. Multiple sequence alignment of the aforementioned proteins with the amino acid sequence from 2-8_MPO showed that these residues are conserved in 2-8_MPO as well. (Figure 4.12).

Even though the structure of 1XAF depicts coordination with zinc ions at the active site, MPOs can harbour different metal ions at the active site. In the case of RL5, for example, the active site contain copper instead of zinc (Beloqui *et al.*, 2006).

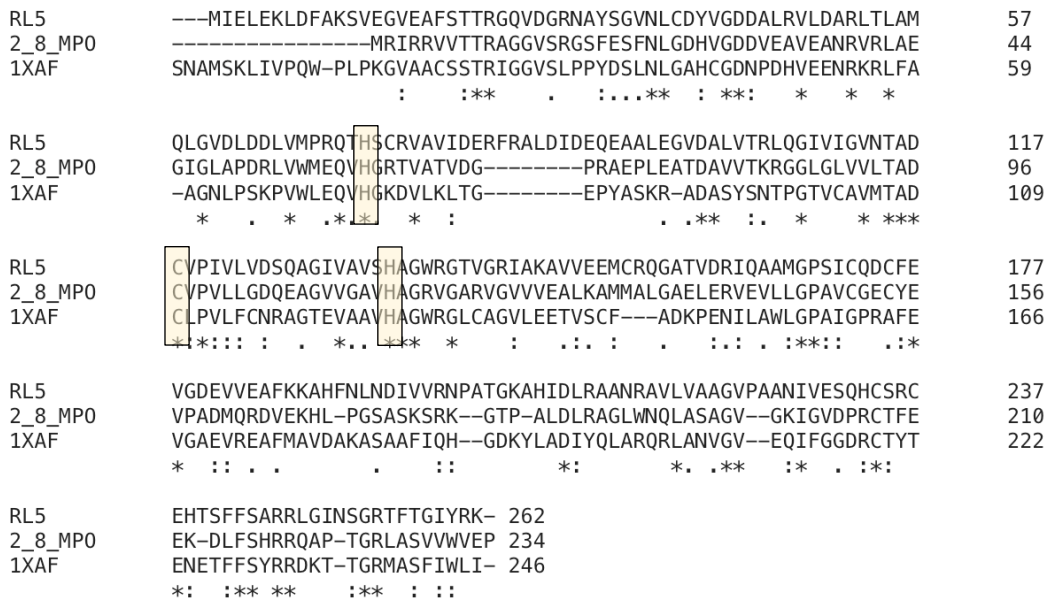


Figure 4.12: Multiple sequence alignment of 2-8_MPO with the YfiH from *S. flexneri* and the RL5 laccase isolated from a metagenomic screen. The highlighted region in the multiple sequence alignment depicts the conserved histidine and cysteine residues thought to be involved in coordinating metal ions in the active site.

MPOs are interesting enzymes due to their extraordinary activity. They often supersede the activities of laccases in terms of the ability to oxidise the 'typical' laccase substrate, such as ABTS, 2,6-DMP and syringaldazine. Beloqui *et al.* (2006) reported that the k_{cat} values for the RL5 laccase were more than 40 times higher than for any laccases previously reported, as well as a redox potential of over 700mV at the active site, which is comparable to high redox potential fungal laccases. This highlights their potential as biocatalysts and necessitates further investigation into what makes them so unique.

4.3.6 Cloning and expressions of MCOs

Ligation reactions for the 2-8_MPO amplicon was unsuccessful. No clones were obtained for this enzyme, despite multiple attempts at reamplifying (which was successful), but no transformants were obtained. The cloning of SF1.4_MCO1 and SF1.4_MCO2 was successful. Confirmation of the desired inserted sizes was confirmed by colony PCR performed on the transformed *E. coli* BL21(DE3) clones (Figure 4.13).

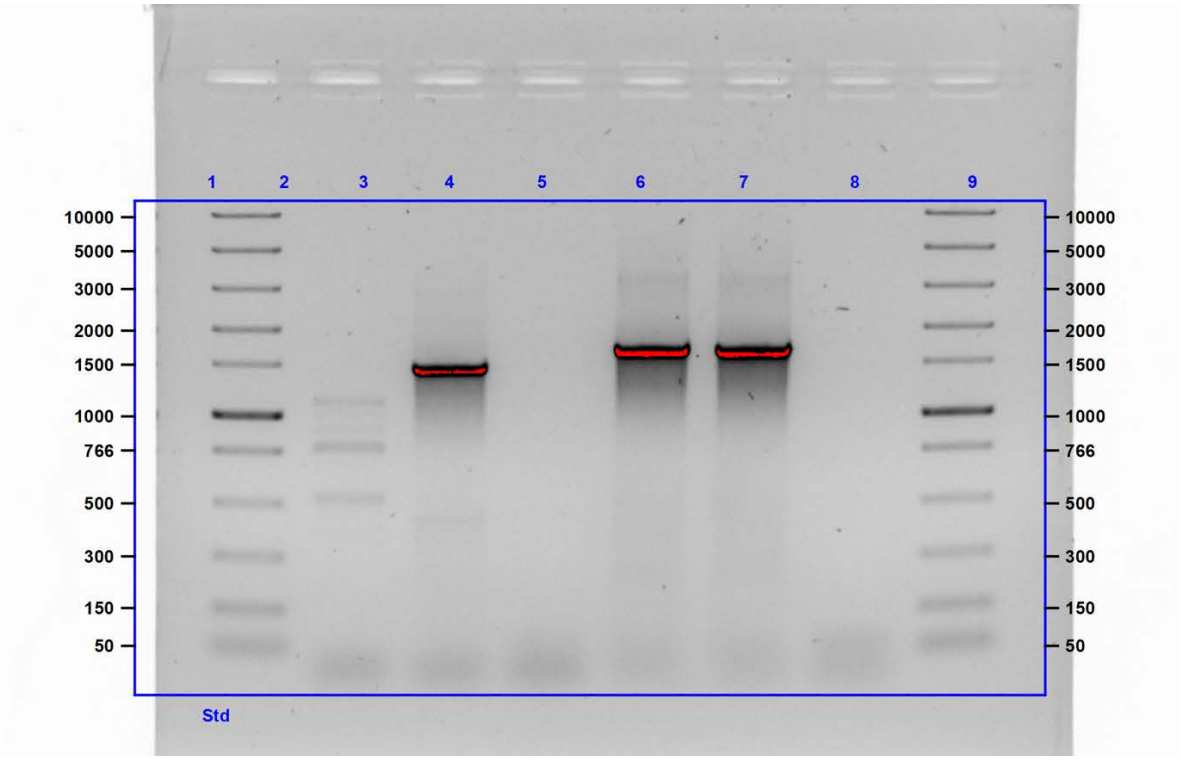


Figure 4.13: Electropherogram of amplicons obtained from colony PCR of *E. coli* BL21(DE3) using the cloning primers for SF1.4_MCO1 (lanes 3 and 4) and SF1.4_MCO2 (lanes 6 and 7).

The constructs were cloned into *E. coli* ArcticExpress RP and *E. coli* BL21(DE3). Two different media types, and two different methods of induction (IPTG and auto-induction) were examined at 15˚C and 22˚C. The ArcticExpress strain was selected to aid in the expression of correctly folded protein – cold-temperature induction has been shown to increase the recovery of soluble protein (Schein & Noteborn, 1988).

Overexpression was mostly unsuccessful for both strains across the different conditions tested – however, low levels of expression was observed for both SF1.4_MCO1 and SF1.4_MCO2 in 2YT media, performing the induction at 22˚C using BL21 as the host (Figure 4.14). Expression was inconsistent between replicate flasks, and the enzymes were inconsistently produced between different expression batches. This inconsistency may likely be due to codon bias in actinobacterial genes. Actinobacteria tend to have a high mol% G+C in their genomes and this is directly linked to the occurrence of rare codons in their genes (Gopal & Kumar, 2013; Lal *et al.*, 2016).

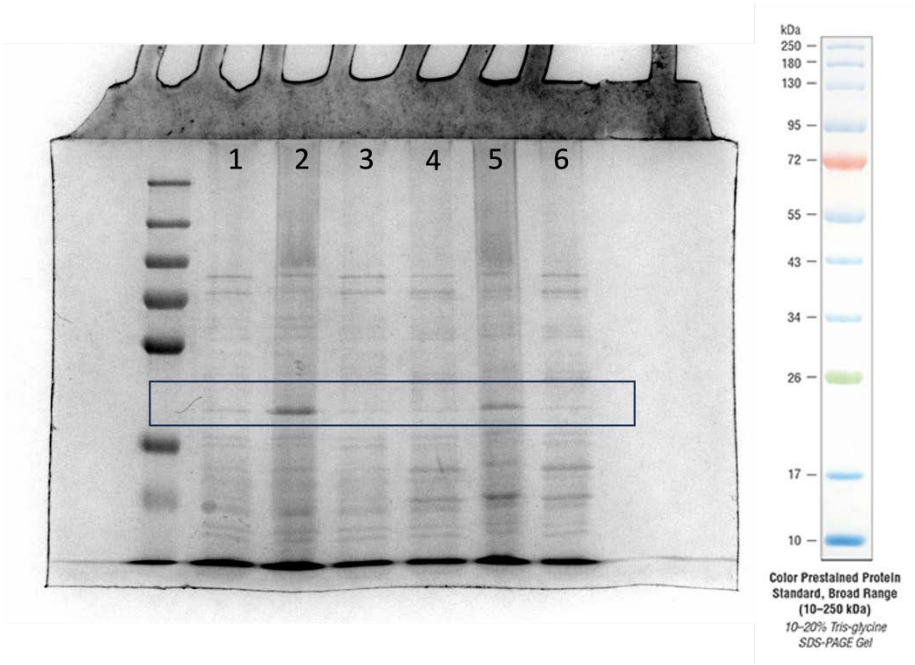


Figure 4.14: SDS-PAGE gel of poorly overexpressed SF1.4_MCO1 (lanes 1-30) and SF1.4_MCO2 (lanes 4-6) in *E. coli* BL21(DE3) using 2YT media, induced at 22˚C. Expected protein size was detected during SDS-PAGE analysis – as highlighted in the boxed area.

As a result of the inconsistent overexpression of the two SF1.4 MCOs, codon-optimised genes were synthesised by SynBio Technologies (New Jersey, United States). The MPO from 2-8 was also included. The codon-optimised constructs were cloned into *E. coli* BL21(DE3) using the process described in section 4.2.6.6.

Clones were not obtained for SF1.4_MCO2. The gene product may possibly have been toxic to cell. However, this is unclear, especially given the similarity between SF1.4_MCO1 and SF1.4_MCO2. The 2-8_MPO and SF1.4_MCO1 constructs, however, successfully transformed and was tested for overexpression in AI media. AI was selected as the medium of choice due to the use of three different carbon sources (including the catabolic repressor, glucose). The growth of *E. coli* is diauxic (exclusively growing on glucose in a medium containing lactose). This has been exploited to increase the density of cell cultures by limiting the amount of glucose (impeding basal expression) in the medium, then forcing lactose utilisation, which in turns activates *lac* operon protein production (Li *et al.*, 2011; Studier, 2005).

This was indeed the case for the 2-8_MPO and SF1.4_MCO1 when using AI media at 30˚C – a definitive increase in production of the two enzymes was observed (Figure 4.15), emphasising the likelihood of codon bias, as well as basal expression, were the causes of non-expression for the original cloned genes. The process was scaled up and a crude enzyme preparation was made to screen for enzyme activity.

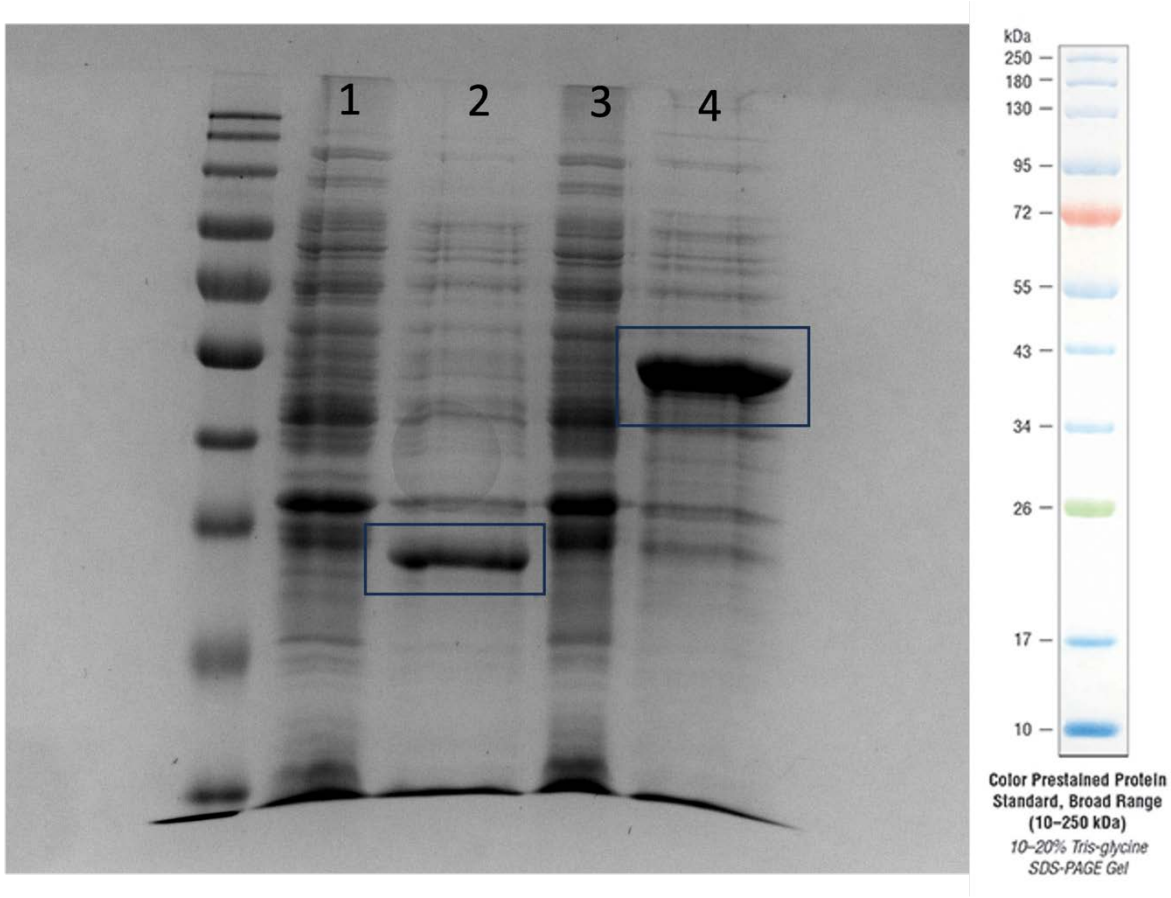


Figure 4.15: SDS-PAGE depicting overexpression of 2-8_MPO (lane 2) and SF1.4_MCO1 (lane 4) produced in *E. coli* BL21(DE3) using AI media, auto-induction taking place at 30˚C. Lanes 1 and 3 depicting the proteins observed for the start cultures.

4.3.7 Preliminary MCO activity measurements

The crude enzyme preparations for 2-8_MPO and SF1.4_MCO were examined for their ability to oxidise 1mM ABTS. First the pH range was determined by monitoring ABTS oxidation between pH 2.0 – pH 5.0 (0.05 mM KH_2PO_4-HCl for pH 2.0 – 2.5; 0.05 mM sodium acetate for pH 3.0 – 5.0). While the measured activity was relatively low (0.07 U/mL for 2-8_MPO and 0.08 U/mL for SF1.4_MCO1), a distinct trend could be observed

with maximal activity exhibited at pH 3.0 and sharply decreasing after pH 3.5 (Figure 4.16). This type of trend is typically observed for MCOs – in this case, comparing 2-8_MPO to the RL5 laccase, and SF1.4_MCO1 to the bilirubin oxidase from *Bacillus pumilis*, a similar trend is noted when ABTS was used as the substrate (Durand *et al.,* 2012; Beloqui *et al.*, 2006).

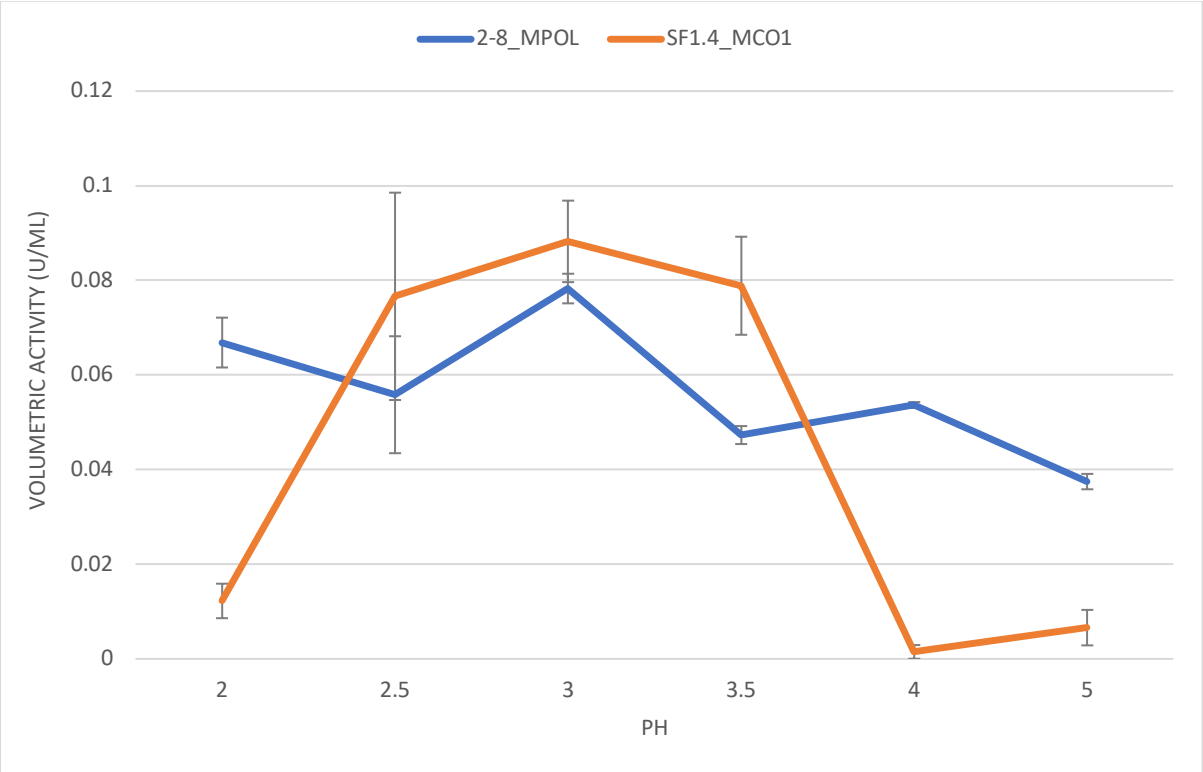


Figure 4.16: The oxidation of ABTS (1mM in 0.05 M KH_2PO_4-HCl [pH 2.0 – pH 2.5], and 0.05 mM sodium acetate for pH 3.0 – 5.0) by 2-8_MPO and SF.14_MCO1. Oxidation was monitored at 420nm. No activity was detected for the BL21 no-plasmid control crude protein extract. Error bars represent the standard deviation observed for triplicate analyses.

The temperature-activity relationship of 2-8_MPO and SF.14_MCO was determined across a temperature range of 25˚C – 45˚C. The activity of 2-8_MPO and SF1.4_MCO1 both increased with an increase in temperature, measuring a maximum of 0.12 U/mL and 0.14 U/mL, respectively. This was unsurprising for 2-8_MPO, when comparing it to the RL5 laccase, which exhibited an increase in activity up to 60˚C (Beloqui *et al.,* 2006). SF1.4_MCO1 showed a similar trend, and while there are no clear reports of temperature-activity dependence in literature, bilirubin oxidases have been shown to be highly thermostable (Durand *et al.*, 2012; Gounel *et al.,* 2016).

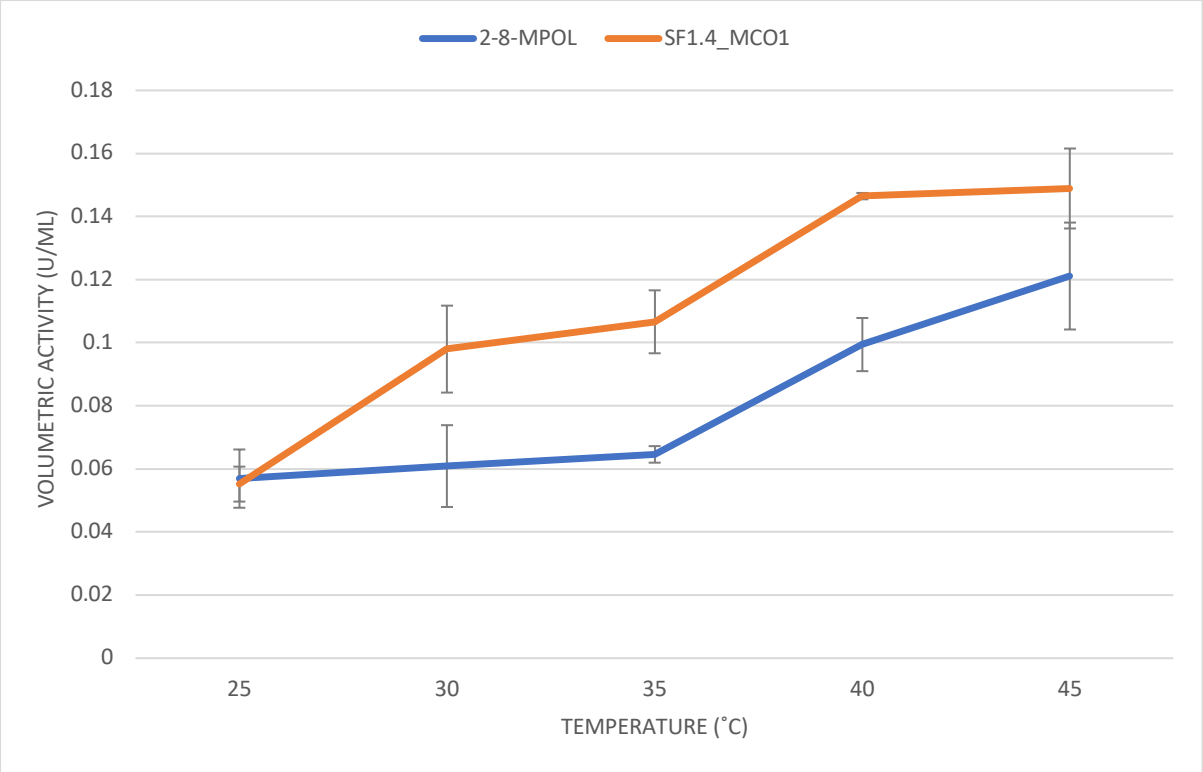


Figure 4.17: Activity-temperature relationship of the ABTS (1mM in 0.05 M KH_2PO_4-HCl [pH 2.0 – pH 2.5], and 0.05 mM sodium acetate for pH 3.0 – 5.0) oxidation activity exhibited by 2-8_MPO and SF.14_MCO1 across a temperature range of 25˚C to 45˚C. Error bars represent the standard deviation observed for triplicate analyses.

4.4 CLOSING REMARKS

Actinobacteria, though extensively studied, can still be considered an enormous source of biotechnological potential, owing to their ability to thrive under a wide range of environmental conditions. While the rise in sequencing technologies has brought about an exponential increase in the amount of sequence data available, there is still great value in pursuing culture-based techniques to access undiscovered actinobacteria. As shown in this study, employing a combination of selective isolation techniques and high-throughput sequencing, while targeting unique environmental conditions, proves advantageous in order to gain access to rare actinobacteria and their enzyme complement.

CHAPTER FIVE
ENZYMEML: KINETIC PARAMETER ESTIMATION THROUGH A PYTHON-BASED WORKFLOW

5.1 PREAMBLE

Reliable and reproducible experimental data is often lacking and becomes a hinderance when assessing the biochemical data for a given enzyme (Begley & Ioannidis, 2015). A considerable discrepancy in intra- and interlaboratory research is often caused by the lack of important information regarding enzymatic experimental data, including important metadata such as pH and temperature conditions, as well as enzyme and substrate concentrations (Baker & Penny, 2016; Halling *et al.*, 2018). EnzymeML, an open data-exchange format was developed by Range *et al.* (2022) to aid in capturing biochemical data from enzymatic experiments according to the guidelines set out by the Standards for Reporting of Enzymology Data (STRENDA) commission.

As part of an ongoing collaboration with researchers at the University of Stuttgart (Germany), EnzymeML was adapted into a Python-based Jupyter-notebook in collaboration with Haüssler (2023) to estimate the kinetic parameters of the small laccase from *Streptomyces coelicolor* (SLAC). The tool consisted of multiple standalone Python modules: a parser to import spectrophotometric measurements taken during the oxidation of ABTS from a pre-prepared EnzymeML spreadsheet, a separate calibration parser to import ABTS standards across a range of concentrations (0-200 μM), and a parameter estimation tool to determine the kinetic parameters of the enzyme.

Here, version two of this Jupyter-notebook is presented. It encompasses a simplified installation process of the required Python modules in a conda environment. Thereafter, the pre-defined Jupyter-notebook is populated with the experimental conditions – measurement data is directly imported from the spectrophotometer output, calibration is performed using control wells already included in the experimental setup and parameter estimation can occur on-the-fly, with the goal of creating a fast and accurate tool for kinetic parameter estimation that can be integrated into laboratory workflows.

5.2 EXPERIMENTAL PROCEDURE

5.2.1 Dataset and computing environment

The data was generated in September of 2022 (accessible through DaRUS; https://doi.org/10.18419/darus-3337). The dataset at the optimal pH and temperature for SLAC (pH 3.0, 45˚C) was used to compare the outputs between version 1 and 2 of the Jupyter-notebook. The enzyme was diluted to a working concentration of 0.1 U/mL (based on the oxidation of 1 mM ABTS in 0.05 M sodium acetate, pH 3.0; 25˚C) of volumetric activity and the protein concentration determined using a standard Bradford assay (Bradford, 1976). ABTS utilisation was measured in triplicate at 340nm across a substrate range of 0 – 200 µM. Triplicate controls for each concentration were prepared as well as an enzyme control. The plate layout is detailed in Figure 5.1.

	1	2	3	4	5	6	7	8	9	10	11	12
A	0 µM	5 µM	10 µM	15 µM	25 µM	50 µM	75 µM	100 µM	125 µM	150 µM	175 µM	200 µM
B	0 µM	5 µM	10 µM	15 µM	25 µM	50 µM	75 µM	100 µM	125 µM	150 µM	175 µM	200 µM
C	0 µM	5 µM	10 µM	15 µM	25 µM	50 µM	75 µM	100 µM	125 µM	150 µM	175 µM	200 µM
D	0 µM	5 µM	10 µM	15 µM	25 µM	50 µM	75 µM	100 µM	125 µM	150 µM	175 µM	200 µM
E	0 µM	5 µM	10 µM	15 µM	25 µM	50 µM	75 µM	100 µM	125 µM	150 µM	175 µM	200 µM
F	0 µM	5 µM	10 µM	15 µM	25 µM	50 µM	75 µM	100 µM	125 µM	150 µM	175 µM	200 µM
G												
H												

Figure 5.1: A 96-well microtiter plate layout for a typical EnzymeML setup. A range of concentrations of ABTS is dispensed from 0-200 µM (columns 1-12). Blue wells denote experimental wells. Green indicates control wells.

Text headers from the original outputs (produced on a SpectraMax M2 Multimode Reader, Molecular Devices, San Jose, USA) were modified to accommodate the new parser written specifically for the plate reader used in this study, a MultiSkan 1000 (ThermoScientific, Massachusetts, USA). The Jupyter-notebook was run on an Apple MacBook Air with an

Apple Silicon M2 Chip containing 8GB of RAM. A Python (version 3.10) conda environment was created to install the 4 Python modules from GitHub (listed in Appendix A).

5.2.2 Jupyter Workflow

[1] After running an assay, the raw measurements are transferred to a folder containing the Jupyter-notebook.

[2] The Jupyter-notebook is opened in an integrated development editor (IDE). In this case, Visual Studio Code version 1.84.1 (Microsoft) was used.

[3] The modules are imported.

[4] Reaction conditions are set (including file path, buffer and enzyme concentrations, temperature, and pH) (Figure 5.2)

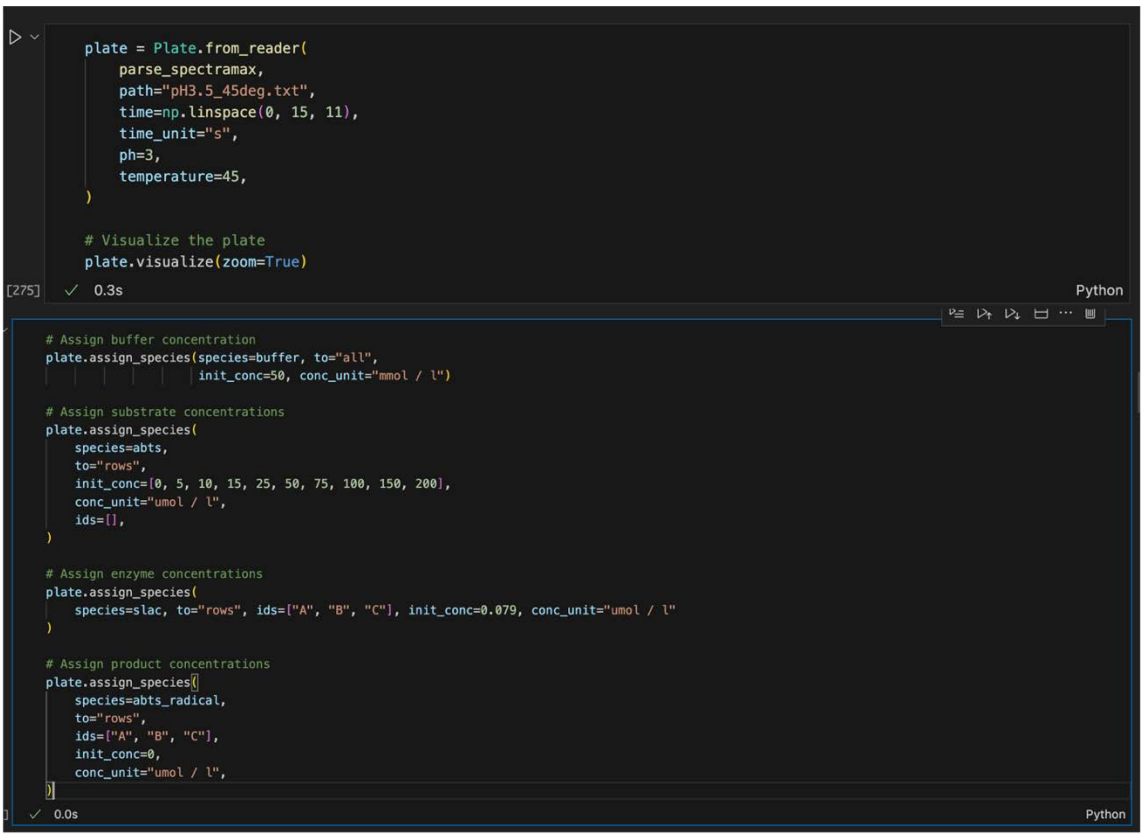


Figure 5.2: Depiction of cells in Jupyter-notebook used to define reaction conditions for the experiment.

[5] Notebook is run, and data output is visualised. A data model with all the relevant information pertaining to the experiment is saved.

The full notebook with all the outputs for each cell is available at the following link (HTML file, viewable in a web browser):
https://github.com/prinsalaric/SLAC_EnzymeML/blob/main/kinetics_SLAC_round1datarevisited.html .

5.3 RESULTS AND DISCUSSION

5.3.1 Benefits

A Jupyter-notebook was designed based on the EnzymeML (Range *et al.*, 2022) data-exchange format to provide a fast and accurate way to determine the kinetic parameters for an enzyme assay. It is meant to provide an end-user friendly Python workflow (regardless of the level of coding expertise). By providing all the relevant metadata and measurements for a single experiment, and the ability to save the outputs in a data model that adheres to the Findable, Accessible, Interoperable and Reproducible (FAIR) principles, it can maximise the usefulness of an individual dataset for real-time use, as well as future research activities (Thomas & Brochu, 2022).

The first version of this EnzymeML Jupyter notebook was designed by Haüssler (2023). It involved several different modules, which required extensive coding knowledge. By incorporating the curation of the raw data measurements and the calibration steps into a singular document, with minimal user input, a rapid workflow can be ascertained. Version 2 was validated using a dataset used to build version 1. The oxidation of ABTS by SLAC was monitored under its optimal conditions (pH 3.0; 45˚C) and it was found that the parameter estimations for k_{cat} and K_m were 0.815 $k_{cat}.s^{-1}$ and 47.96 µM, and 0.804 $k_{cat}.s^{-1}$ and 47.01 µM, for versions 1 and 2, respectively.

The pipeline, however, at this stage is still rather limited. Only ABTS can currently be used as a substrate, and the parser is limited to a single instrument for measurements. By expanding the capabilities of the pipeline, it would allow for better integration into existing workflows, and allow users with an easier way to incorporate relevant information from their experiments.

5.3.2 Concluding remarks

Critical information regarding enzymatic reactions is often lacking which leads to problems with reproducibility (Halling *et al.*, 2018). Reviewing 36 entries for SLAC in SABIO-RK, a database that extracts biochemical data from literature (Wittig *et al.*, 2018), a large discrepancy in reported parameters can be observed (Figure 5.3).

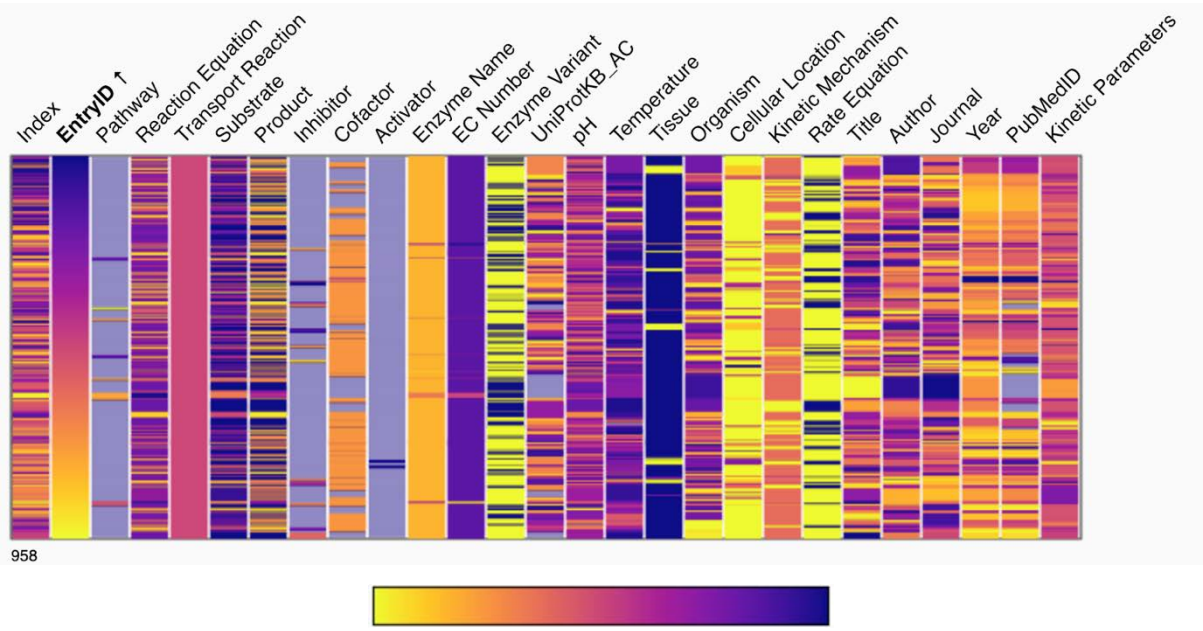


Figure 5.3: Heatmap showing distribution of biochemical parameters for laccases reported from literature on SABIO-RK (accessed 28 October 2023). Lighter colour = less frequently reported/absence of information; dark colour = more frequently reported / detail provided.

Halling *et al.* (2018) did an empirical review of biochemical data presented in articles and highlighted missing information. These range from poorly defined reactions (stating kinetic parameters but not providing sufficient time-course data, or any at all), values lacking units to presenting certain findings in text, but contradictory supplementary information is provided. The proper reporting of enzymology data is imperative for enabling reproducibility in the field. By employing tools such as EnzymeML to ascertain that "good" data is generated, reliable and open science can take place.

CHAPTER SIX
GENERAL CONCLUSIONS

6.1 SUMMARY OF STUDY FINDINGS

The Cape Floristic Region (CFR) is one of the most biodiverse biomes in South Africa with a unique plant biodiversity (Rutherford *et al.,* 2006). It is a region boasting a plant endemicity of almost 75% (Goldblatt, 1978). The Fynbos biome, which forms part of the Greater CFR, can be characterised by the prevalence of low to medium height shrubland, including 81 distinct true and fire-prone fynbos vegetation units, and 38 renosterveld and sandveld units (Low & Rebelo, 1998). Plant species are typically found in soils that are acid and nutrient-poor (Manning, 2018). Understanding what drives these unique environmental conditions ranging from low-nutrient to organic-rich ecosystems, may also provide information regarding the microbial diversity. A limited number of studies report the actinobacterial diversity associated with the CFR. In-depth analyses of the CFR may provide insights into these organisms and whether they produce secondary metabolites of biotechnological interest.

6.1.1 Actinobacterial diversity associated with the CFR

Microorganisms are key players in biogeochemical cycles – plant-soil microbe interactions are imperative for maintaining healthy soil and plant ecosystems (Le Roux *et al.*, 2017). Actinobacteria contribute a significant portion of key indicator species that have been reported to be present in soil communities (Pankratov *et al.*, 2006). To determine the actinobacterial complement associated with two CFR regions (soils associated with Rooibos in Clanwilliam, and soils associated with an emerging peatland in Springfield, Agulhas), 16S rRNA-based metabarcoding was performed. The top five major taxa contributing to the Rooibos environment were identified as *Mycobacteriaceae, Pseudonocardiaceae, Frankiaceae* and *Geodermatophilaceae*. Less prominent taxa include members of the families *Microbacteriaceae, Micromonosporaceae, Nocardiodaceae*, *Acidimicrobiaceae* and *Nakamurellaceae*. For the Springfield environment, it was found that it was predominated by members of the families *Mycobacteriaceae,* koll13_f &

AKIW874_f (class *Acidimicrobiia*), *Pseudonocardiaceae* and *Nocardioidaceae*. Minor taxa include *Micromonosporaceae, Geodermatophilaceae, Nocardiopsceae* and *Intrasporangiaceae*.

Even though an overall snapshot of the actinobiome associated with the two CFR regions was determined, several limitations apply to this study. Firstly, the use of actinobacterial-specific primers may have introduced bias by preferentially amplifying specific taxa (Acinas *et al.*, 2003). To overcome this bias in future studies, shotgun sequencing of the entire metagenome is recommended, together with up-to-date sequence data, in order to elucidate the microbial communities at a higher level of accuracy (Stach *et al.*, 2003; van der Walt *et al.*, 2017). Furthermore, while not in the scope of this study, a single sampling time point is simply not enough to infer soil-plant-microbe interactions. This may be addressed by monitoring various environmental conditions, as well as the microbial communities, over longer periods of time to gain insights into what drives the unique conditions observed for the CFR.

6.1.2 Rare actinobacteria from the explored CFR regions

Actinobacteria are Gram-positive bacteria that are widely distributed in nature (Macagnan *et al.,* 2006). These organisms have the ability to survive under a number of environmental conditions including but not limited to, extremes of temperatures, acidity, alkalinity and nutrient poor environments (such as the CFR). Actinobacteria are known to produce a wide range of industrially relevant secondary metabolites, including enzymes such as multicopper oxidases (MCOs) (Goodfellow *et al.,* 2018; Shivlata & Satyanarana, 2015). Rare actinobacteria are of particular interest, as not much is known regarding the MCOs that are produced by these underrepresented organisms. Cross-referencing the taxa identified in this study with entries in the Laccase and Multicopper Oxidase Engineering Database (LccED), only a select number of the major genera identified in the Rooibos and Springfield environments were represented. This presented a unique opportunity to explore these environments further for the presence of MCOs from rare actinobacteria through selective isolation techniques.

The selective isolation of actinobacteria from Rooibos soil was performed using seven different isolation media that were supplemented with penicillin and cycloheximide to inhibit the growth of non-actinobacteria and fungi. The choice of selective media, including a medium containing Rooibos extract, along with pre-treatments was guided by methods outlined by Hayakawa *et al.* (2008) for the isolation of rare actinobacteria. The isolates were screened for the presence of extracellular MCO activity using ABTS and 2,6-dimethoxyphenol (DMP) as substrates. The top 25 MCO-producing isolates were identified through 16S rRNA gene sequence analyses. Eight out of the 25 were identified as rare actinobacteria, including members of the genera *Dactylosporangium, Actinokineospora, Curtobacterium, Modestobacter, Leifsonia* and *Actinomadura.* The genera are vastly underrepresented in the LccED, with only 74 records obtained among a total of 3 400 bacterial MCOs represented across five superfamilies.

6.1.3 MCOs isolated from rare actinobacteria

MCOs are enzymes of industrial relevance owing to their wide range of activities. This means they can be applied in the several applications such as the degradation of lignocellulosic materials and the detoxification of environmental pollutants (Le Roes-Hill and Prins, 2016). Through genome mining of the top 5 strains, thirteen potentially new copper oxidases were identified – 9 putative multicopper oxidases and 4 multicopper polyphenol oxidases (MPOs). Based on selection criteria (which included homology to known domain architectures, putative catalytic sites, and predicted activity), two MCOs were selected designated SF1.4_MCO1 and SF1.4_MCO2 containing domain architecture that were typical of three-domain MCOs (3dMCOs). A third, rare MPO (designated 2-8_MPO) was also identified showing homology to YfiH and the RL5 laccase.

Functional expression was obtained through heterologous expression. The initial stumbling blocks highlighted the need to carefully approach the heterologous expression of actinobacterial genes – this is considerably challenging due to a number of factors, including rare codon bias. While generally acceptable for proteins under 100 kDa, the occurrence of even a single rare codon has been shown to influence overexpression of genes from organisms that possess high mol% G+C content in their DNA (Lal *et al.,* 2016). Nonetheless, 2-8_MPO and SF1.4_MCO1 was sufficiently expressed when codon-optimised constructs were synthesised and expressed in *E. coli* BL21(DE3), and the crude

preparations were used to obtain some baseline biochemical data. Both 2-8_MPO and SF1.4_MCO1 exhibited activity at pH 3.0 using ABTS as the substrate. Maximal activity for each enzyme was observed at 45˚C. This was particularly interesting in the case of 2-8_MPO. MPOs are remarkable enzymes as they often surpass the activities of known laccases by several orders of magnitude, making them interesting targets for biotechnological development. One such MPO is the RL5 that was cloned from a metagenome screening study. Beloqui *et al.* (2006) has reported k_{cat} values greater than any other laccase before, despite not adhering to the classic features of an MCO – which highlights another issue. Often laccases are assigned their identity based solely on sequence similarity or the ability to oxidise a specific substrate. While this is useful, as observed in this study using protein databases such as the LccED for discovering new MCOs, a more holistic approach should be taken in order not to miss out on potential enzymes that may be useful in biocatalytic processes.

6.2 RECOMMENDATIONS FOR FUTURE STUDIES

This is the first reports of MCOs from these two unique South African environments, especially from rare actinobacteria. It was demonstrated here that by combining selective isolation techniques, culture-based screening, and genome mining, unexplored actinobacteria are a potential source of new enzymes that may exhibit novel activities or properties.

While the initial experiments determined that the heterologously expressed enzymes were functional, optimisation is required to increase yield. This may be achieved through the optimisation of culture conditions (aeration, optical density selected when inducing overexpression, varying induction temperatures), the use of different *E. coli* hosts, or even expression in non-*E. coli* hosts such as the yeast *Pichia pastoris* or an actinobacterium such as *Rhodococcus erythropolis.* Furthermore, purification strategies need to be investigated in order to determine the kinetic parameters of the new enzymes using a platform such as EnzymeML. Currently, the use of EnzymeML is limited to the monitoring of ABTS substrate utilisation. The capabilities of the workflow should also be extended to include more of the ‘typical’ MCO substrates in order to provide a detailed workflow for kinetic parameter estimation, within the guidelines described by the STRENDA

commission for the generation of high-quality biochemical data for 2-8_MPO and SF1.4_MCO1.

Currently, there is limited information available on MPOs, particularly for extensive biochemical data and structural elucidation. One strategy to address this is the use of sequence networks to study the links between these atypical enzymes to others. To address this problem, the Python Enzyme Engineering Database (PyEED) is currently under development, in collaboration with the University of Stuttgart. The aim of this project is to not only generate real-time enzyme databases by incorporating new sequences as they are released, but also to incorporate available enzymology data to create sequence networks which would allow to gain more insights into how these enzymes evolve, and what the key drivers are that confer certain activity profiles. Using a tool such as PyEED would be useful in the further exploration of these new MPOs, or any other enzymes for that matter.

REFERENCES

Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., & Polz, M. F. (2005). PCR-Induced Sequence Artifacts and Bias: Insights from Comparison of Two 16S rRNA Clone Libraries Constructed from the Same Sample. *Applied and Environmental Microbiology*, *71*(12), 8966–8969. https://doi.org/10.1128/AEM.71.12.8966-8969.2005

Ahmed, I. H., Francina, L. B., Isabella, H. R., & Galaletsang, S. (2014). Nodulation efficacy of Bradyrhizobium japonicum inoculant strain WB74 on soybean (Glycine max L. Merrill) is affected by several limiting factors. *African Journal of Microbiology Research*, *8*(20). https://doi.org/10.5897/ajmr2014.6709

Aksenov, A. S., Shirokova, L. S., Kisil, O. Y., Kolesova, S. N., Lim, A. G., Kuzmina, D., Pouillé, S., Alexis, M. A., Castrec-Rouelle, M., Loiko, S. V., & Pokrovsky, O. S. (2021). Bacterial number and genetic diversity in a permafrost peatland (Western siberia): Testing a link with organic matter quality and elementary composition of a peat soil profile. *Diversity*, *13*(7). https://doi.org/10.3390/d13070328

Anandan, R., Dharumadurai, D., & Manogaran, G. P. (2016). 'An Introduction to Actinobacteria', in Dharumadurai Dhanasekaran and Yi Jiang (eds) *Actinobacteria*. Rijeka: IntechOpen, p. Ch. 1. Available at: https://doi.org/10.5772/62329.

Alcalde, M. (2007). Laccases: Biological functions, molecular structure and industrial applications. In *Industrial Enzymes: Structure, Function and Applications*. https://doi.org/10.1007/1-4020-5377-0_26

Alves, T., Fernandes, R., Batista, W., Maria, F., Passos, L., Zucchi, T. D., Fernandes, T. A. R., Silveira, W. B. da, & Passos, F. M. L. (2014). Laccases from Actinobacteria—What We Have and What to Expect. *Advances in Microbiology*, *4*(May).

Arkin, A. P., Cottingham, R. W., Henry, C. S., Harris, N. L., Stevens, R. L., Maslov, S., Dehal, P., Ware, D., Perez, F., Canon, S., Sneddon, M. W., Henderson, M. L., Riehl, W. J., Murphy-Olson, D., Chan, S. Y., Kamimura, R. T., Kumari, S., Drake, M. M., Brettin, T. S., … Yu, D. (2018). KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nature Biotechnology*, *36*(7), 566–569. https://doi.org/10.1038/nbt.4163

Aziz, R. K., Bartels, D., Best, A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., … Zagnitko, O. (2008). The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics*, *9*. https://doi.org/10.1186/1471-2164-9-75

Baba, H., Ohtsuka, Y., Haruna, H., Lee, T., Nagata, S., Maeda, M., Yamashiro, Y., & Shimizu, T. (2009). Studies of anti-inflammatory effects of Rooibos tea in rats. *Pediatrics International*, *51*(5), 700–704.

Baker, M., & Penny, D. (2016). Is there a reproducibility crisis? In *Nature* (Vol. 533, Issue 7604). https://doi.org/10.1038/533452A

Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H.-P., Clément, C., Ouhdouch, Y., & Wezel, G. P. van. (2016). Taxonomy, physiology, and natural products of Actinobacteria. *Microbiology and Molecular Biology Reviews*, *80*(1), 1–43.

Bateman, A., Martin, M. J., Orchard, S., Magrane, M., Ahmad, S., Alpi, E., Bowler-Barnett, E. H., Britto, R., Bye-A-Jee, H., Cukura, A., Denny, P., Dogan, T., Ebenezer, T. G., Fan, J., Garmiri, P., Gonzales, L. J. da C., Hatton-Ellis, E., Hussein, A., Ignatchenko, A., … Zhang, J. (2023). UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*, *51*(D1). https://doi.org/10.1093/nar/gkac1052

Begley, C. G., & Ioannidis, J. P. A. (2015). Reproducibility in science: Improving the standard for basic and preclinical research. In *Circulation Research* (Vol. 116, Issue 1). https://doi.org/10.1161/CIRCRESAHA.114.303819

Behr, A. S., Abbaspour, E., Rosenthal, K., Pleiss, J., & Kockmann, N. (2023). Ontology-Based Laboratory Data Acquisition With EnzymeML for Process Simulation of Biocatalytic Reactors. *Proceedings of the Conference on Research Data Infrastructure*, *1*. https://doi.org/10.52825/cordi.v1i.324

Beloqui, A., Pita, M., Polaina, J., Martínez-Arias, A., Golyshina, O. V., Zumárraga, M., Yakimov, M. M., García-Arellano, H., Alcalde, M., Fernández, V. M., Elborough, K., Andreu, J. M., Ballesteros, A., Plou, F. J., Timmis, K. N., Ferrer, M., & Golyshin, P. N. (2006). Novel polyphenol oxidase mined from a metagenome expression library of bovine rumen: Biochemical properties, structural analysis, and phylogenetic relationships. *Journal of Biological Chemistry*, *281*(32). https://doi.org/10.1074/jbc.M600577200

Béna, G., Lyet, A., Huguet, T., & Olivieri, I. (2005). Medicago—Sinorhizobium symbiotic specificity evolution and the geographic expansion of Medicago. *Journal of Evolutionary Biology*, *18*(6). https://doi.org/10.1111/j.1420-9101.2005.00952.x

Bennett, J. A., Kandell, G. V., Kirk, S. G., & McCormick, J. R. (2018). Visual and Microscopic Evaluation of Streptomyces Developmental Mutants. *Journal of Visualized Experiments*, *139*, 57373. https://doi.org/10.3791/57373

Bérdy, J. (2005). Bioactive Microbial Metabolites. *The Journal of Antibiotics*, *58*(1). https://doi.org/10.1038/ja.2005.1

Bergmann, F. T., Hoops, S., Klahn, B., Kummer, U., Mendes, P., Pahle, J., & Sahle, S. (2017). COPASI and its applications in biotechnology. In *Journal of Biotechnology* (Vol. 261). https://doi.org/10.1016/j.jbiotec.2017.06.1200

Born, J., Linder, H. P., & Desmet, P. (2007). The Greater Cape Floristic Region. *Journal of Biogeography*, *34*(1). https://doi.org/10.1111/j.1365-2699.2006.01595.x

Botta, A. (2012). Enhancing plant tolerance to temperature stress with amino acids: An approach to their mode of action. *Acta Horticulturae*, *1009*. https://doi.org/10.17660/actahortic.2013.1009.1

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, *72*(1–2), 248–254. https://doi.org/10.1016/0003-2697(76)90527-3

Bredholdt, H., Galatenko, O. A., Engelhardt, K., Fjærvik, E., Terekhova, L. P., & Zotchev, S. B. (2007). Rare actinomycete bacteria from the shallow water sediments of the Trondheim fjord, Norway: Isolation, diversity and biological activity. *Environmental Microbiology*, *9*(11). https://doi.org/10.1111/j.1462-2920.2007.01387.x

Brink, C. J., Postma, A., Slabbert, E., Postma, F., Muasya, A. M., & Jacobs, K. (2020). Bacterial communities associated with natural and commercially grown rooibos (Aspalathus linearis). *Pedosphere*, *30*(6). https://doi.org/10.1016/S1002-0160(20)60038-8

Burton, S. (2005). Laccases and Phenol Oxidases in Organic Synthesis—A Review. *Current Organic Chemistry*, *7*(13). https://doi.org/10.2174/1385272033486477

Cabana, H., Jones, J. P., & Agathos, S. N. (2009). Utilization of cross-linked laccase aggregates in a perfusion basket reactor for the continuous elimination of endocrine-disrupting chemicals. *Biotechnology and Bioengineering*, *102*(6). https://doi.org/10.1002/bit.22198

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Pẽa, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., … Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. In *Nature Methods* (Vol. 7, Issue 5). https://doi.org/10.1038/nmeth.f.303

Chai, B., Qiao, Y., Wang, H., Zhang, X., Wang, J., Wang, C., Zhou, P., & Chen, X. (2017). Identification of YfiH and the catalase catA as polyphenol oxidases of aeromonas media and catA as a regulator of pigmentation by its peroxyl radical scavenging capacity. *Frontiers in Microbiology*, *8*(OCT). https://doi.org/10.3389/fmicb.2017.01939

Chang, A., Jeske, L., Ulbrich, S., Hofmann, J., Koblitz, J., Schomburg, I., Neumann-Schaal, M., Jahn, D., & Schomburg, D. (2021). BRENDA, the ELIXIR core data resource in 2021: New developments and updates. *Nucleic Acids Research*, *49*(D1). https://doi.org/10.1093/nar/gkaa1025

Claus, H. (2004). Laccases: Structure, reactions, distribution. *Micron*, *35*(1–2). https://doi.org/10.1016/j.micron.2003.10.029

Cook, A. E. (2003). Rapid identification of filamentous actinomycetes to the genus level using genus-specific 16S rRNA gene restriction fragment patterns. *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY*, *53*(6), 1907–1915. https://doi.org/10.1099/ijs.0.02680-0

Cowan, G. I. (1995). Wetland regions of South Africa. *Wetlands of South Africa*, 21–31.

Cowling, R. M. (1990). Diversity components in a species-rich area of the Cape Floristic Region. *Journal of Vegetation Science*, 699–710.

Dahlgren, R. (1968). REVISION OF GENUS ASPALATHUS. 2. SPECIES WITH ERICOID AND PINOID LEAFLETS. 7. SUBGENUS NORTIERIA. WITH REMARKS ON ROOIBOS TEA CULTIVATION. *Botaniska Notiser*, *121*(2), 165.

DEA. (2014). *Traditional Knowledge Associated with Rooibos and Honeybush Species in South Africa*. Department of Environmental Affairs (DEA). https://www.environment.gov.za/sites/default/files/reports/traditionalknowledge_rooibosandhoneybushspecies_report.pdf

Dowd, S. E., Wolcott, R. D., Sun, Y., McKeehan, T., Smith, E., & Rhoads, D. (2008). Polymicrobial Nature of Chronic Diabetic Foot Ulcer Biofilm Infections Determined Using Bacterial Tag Encoded FLX Amplicon Pyrosequencing (bTEFAP). *PLoS ONE*, *3*(10), e3326. https://doi.org/10.1371/journal.pone.0003326

Driver, A., Sink, K. J., Nel, J. L., Holness, S., Niekerk, L. V., Daniels, F., Jonas, Z., Majiedt, P. A., Harris, L., & Maze, K. (2012). *National Biodiversity Assessment 2011: An assessment of South Africa's biodiversity and ecosystems.*

Durand, F., Kjaergaard, C. H., Suraniti, E., Gounel, S., Hadt, R. G., Solomon, E. I., & Mano, N. (2012). Bilirubin oxidase from Bacillus pumilus: A promising enzyme for the elaboration of efficient cathodes in biofuel cells. *Biosensors and Bioelectronics*, *35*(1), 140–146. https://doi.org/10.1016/j.bios.2012.02.033

Dwivedi, U. N., Singh, P., Pandey, V. P., & Kumar, A. (2011). Structure-function relationship among bacterial, fungal and plant laccases. In *Journal of Molecular Catalysis B: Enzymatic* (Vol. 68, Issue 2). https://doi.org/10.1016/j.molcatb.2010.11.002

Eggert, C., Temp, U., & Eriksson, K. E. (1996). The ligninolytic system of the white rot fungus Pycnoporus cinnabarinus: Purification and characterization of the laccase. *Applied and Environmental Microbiology*, *62*(4), 1151–1158. https://doi.org/10.1128/aem.62.4.1151-1158.1996

Ellery, W. N., Grenfell, S. E., Grenfell, M. C., Humphries, M. S., Barnes, K., Dahlberg, A., & Kindness, A. (2012). Peat formation in the context of the development of the Mkuze floodplain on the coastal plain of Maputaland, South Africa. *Geomorphology*, *141*, 11–20.

Enguita, F. J., Martins, L. O., Henriques, A. O., & Carrondo, M. A. (2003). Crystal Structure of a Bacterial Endospore Coat Component. *Journal of Biological Chemistry*, *278*(21). https://doi.org/10.1074/jbc.m301251200

Erickson, L. (2003). Rooibos tea: Research into antioxidant and antimutagenic properties. *HerbalGram*, *59*, 34–45.

Falkinham, J. O. (2009). The biology of environmental mycobacteria. In *Environmental Microbiology Reports* (Vol. 1, Issue 6). https://doi.org/10.1111/j.1758-2229.2009.00054.x

Fang, B.-Z., Salam, N., Han, M.-X., Jiao, J.-Y., Cheng, J., Wei, D.-Q., Xiao, M., & Li, W.-J. (2017). Insights on the Effects of Heat Pretreatment, pH, and Calcium Salts on Isolation of Rare Actinobacteria from Karstic Caves. *Frontiers in Microbiology*, *8*, 1535. https://doi.org/10.3389/fmicb.2017.01535

Fekry, R. M., Ezz, M. K., Keshta, A. T., & Hazem, W. M. (2014). Cardio-Protective Effect of Rooibos against Isoproterenol Induced Myocardial Infarction in Male Albino Rats. *Biochemistry Letters*, *9*(1). https://doi.org/10.21608/blj.2014.63854

Franche, C., Lindström, K., & Elmerich, C. (2009). Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants. *Plant and Soil*, *321*(1–2). https://doi.org/10.1007/s11104-008-9833-8

Frolking, S., Roulet, N., & Fuglestvedt, J. (2006). How northern peatlands influence the Earth's radiative budget: Sustained methane emission versus sustained carbon sequestration. *Journal of Geophysical Research: Biogeosciences*, *111*(G1).

Giardina, P., Faraco, V., Pezzella, C., Piscitelli, A., Vanhulle, S., & Sannia, G. (2010). Laccases: A never-ending story. *Cellular and Molecular Life Sciences*, *67*(3). https://doi.org/10.1007/s00018-009-0169-1

Gihaz, S., Herzallh, N. S., Cohen, Y., Bachar, O., Fishman, A., & Yehezkeli, O. (2022). The Structure of Bilirubin Oxidase from Bacillus pumilus Reveals a Unique Disulfide Bond for Site-Specific Direct Electron Transfer. *Biosensors*, *12*(5), 258. https://doi.org/10.3390/bios12050258

Gilbert, D., & Mitchell, E. A. D. (2006). Microbial diversity in Sphagnum peatlands. *Developments in Earth Surface Processes*, *9*, 287–318.

Gillespie, J. J., Wattam, A. R., Cammer, S. A., Gabbard, J. L., Shukla, M. P., Dalay, O., Driscoll, T., Hix, D., Mane, S. P., Mao, C., Nordberg, E. K., Scott, M., Schulman, J. R., Snyder, E. E., Sullivan, D. E., Wang, C., Warren, A., Williams, K. P., Xue, T., … Sobral, B. W. (2011). PATRIC: The Comprehensive Bacterial Bioinformatics Resource with a Focus on Human Pathogenic Species. *Infection and Immunity*, *79*(11), 4286–4298. https://doi.org/10.1128/IAI.00207-11

Glickman, C. M., Virdi, R., Hasan, N. A., Epperson, L. E., Brown, L., Dawrs, S. N., Crooks, J. L., Chan, E. D., Strong, M., Nelson, S. T., & Honda, J. R. (2020). Assessment of Soil Features on the Growth of Environmental Nontuberculous Mycobacterial Isolates from Hawai'i. *Applied and Environmental Microbiology*, *86*(21). https://doi.org/10.1128/AEM.00121-20

Goldblatt, P. (1978). An analysis of the flora of southern Africa: Its characteristics, relationships, and orgins. *Annals of the Missouri Botanical Garden*, 369–436.

Gonzalo, G. de, & Franconetti, A. (2018). Enantioselective sulfoxidations employing the thermostable cyclohexanone monooxygenase from Thermocrispum municipale. *Enzyme and Microbial Technology*, *113*. https://doi.org/10.1016/j.enzmictec.2018.02.006

Gonzalo, G. D., Rodríguez, C., Rioz-Martínez, A., & Gotor, V. (2012). Improvement of the biocatalytic properties of one phenylacetone monooxygenase mutant in hydrophilic organic solvents. *Enzyme and Microbial Technology*, *50*(1). https://doi.org/10.1016/j.enzmictec.2011.09.006

Goodfellow, M., Nouioui, I., Sanderson, R., Xie, F., & Bull, A. T. (2018). Rare taxa and dark microbial matter: Novel bioactive actinobacteria abound in Atacama Desert soils. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *111*(8). https://doi.org/10.1007/s10482-018-1088-7

Gopal, G. J., & Kumar, A. (2013). Strategies for the Production of Recombinant Protein in Escherichia coli. *The Protein Journal*, *32*(6), 419–425. https://doi.org/10.1007/s10930-013-9502-5

Gounel, S., Rouhana, J., Stines-Chaumeil, C., Cadet, M., & Mano, N. (2016). Increasing the catalytic activity of Bilirubin oxidase from Bacillus pumilus: Importance of host strain and chaperones proteins. *Journal of Biotechnology*, *230*, 19–25. https://doi.org/10.1016/j.jbiotec.2016.04.035

Grace, J. (2004). Understanding and managing the global carbon cycle. *Journal of Ecology*, *92*(2). https://doi.org/10.1111/j.0022-0477.2004.00874.x

Grace, J., Mitchard, E., & Gloor, E. (2014). Perturbations in the carbon budget of the tropics. *Global Change Biology*, *20*(10), 3238–3255. https://doi.org/10.1111/gcb.12600

Gräff, M., Buchholz, P. C. F., Roes-Hill, M. L., & Pleiss, J. (2020). Multicopper oxidases: Modular structure, sequence space, and evolutionary relationships. *Proteins: Structure, Function and Bioinformatics*, *88*(10). https://doi.org/10.1002/prot.25952

Grundling, P., Grundling, A. T., Pretorius, L., Mulders, J., & Mitchell, S. (2017). South African peatlands: Ecohydrological characteristics and socio-economic value. In *Water Research Commission: Pretoria, South Africa*.

Gründling, P.-L., & Grobler, R. (2005). Peatlands and Mires of South Africa. *Stapfia*, *35*.

Gunne, M., & Urlacher, V. B. (2012). Characterization of the Alkaline Laccase Ssl1 from Streptomyces sviceus with Unusual Properties Discovered by Genome Mining. *PLoS ONE*, *7*(12). https://doi.org/10.1371/journal.pone.0052360

Halling, P., Fitzpatrick, P. F., Raushel, F. M., Rohwer, J., Schnell, S., Wittig, U., Wohlgemuth, R., & Kettner, C. (2018). An empirical analysis of enzyme function reporting for experimental reproducibility: Missing/incomplete information in published papers. *Biophysical Chemistry*, *242*. https://doi.org/10.1016/j.bpc.2018.08.004

Hassen, A. I., Habig, J. H., & Lamprecht, S. C. (2018). Assessing root nodule microsymbionts in healthy and declined rooibos (Aspalathus linearis burm f.) at a plantation in south africa. *Journal of Plant Interactions*, *13*(1). https://doi.org/10.1080/17429145.2018.1473514

Häußler, M. (2023). *From raw data to kinetic parameters: An EnzymeML-based workflow for reproducible enzyme kinetics* [dataset]. DaRUS. https://doi.org/10.18419/DARUS-3337

Hayakawa, M. (2008). Studies on the Isolation and Distribution of Rare Actinomycetes in Soil. *Actinomycetologica*, *22*(1). https://doi.org/10.3209/saj.saj220103

Hayakawa, M., & Nonomura, H. (1987). Efficacy of artificial humic acid as a selective nutrient in HV agar used for the isolation of soil actinomycetes. *Journal of Fermentation Technology*, *65*(6), 609–616. https://doi.org/10.1016/0385-6380(87)90001-X

Heerden, F. R. V., Wyk, B. E. V., Viljoen, A. M., & Steenkamp, P. A. (2003). Phenolic variation in wild populations of Aspalathus linearis (rooibos tea). *Biochemical Systematics and Ecology*, *31*(8). https://doi.org/10.1016/S0305-1978(03)00084-X

Hendey, Q. B. (1983). Palaeontology and palaeoecology of the Fynbos region: An introduction. In *Fynbos palaeoecology: A preliminary synthesis. South African National Scientific Programmes Report* (Vol. 75, pp. 87–99).

Herrera-Zúñiga, L. D., Millán-Pacheco, C., Viniegra-González, G., Villegas, E., Arregui, L., & Rojo-Domínguez, A. (2019). Molecular dynamics on laccase from Trametes versicolor to examine thermal stability induced by salt bridges. *Chemical Physics*, *517*. https://doi.org/10.1016/j.chemphys.2018.10.019

Hon, J., Marusiak, M., Martinek, T., Kunka, A., Zendulka, J., Bednar, D., & Damborsky, J. (2021). SoluProt: Prediction of soluble protein expression in *Escherichia coli*. *Bioinformatics*, *37*(1), 23–28. https://doi.org/10.1093/bioinformatics/btaa1102

Hong, K., Gao, A. H., Xie, Q. Y., Gao, H., Zhuang, L., Lin, H. P., Yu, H. P., Li, J., Yao, X. S., Goodfellow, M., & Ruan, J. S. (2009). Actinomycetes for marine drug discovery isolated from mangrove soils and plants in China. *Marine Drugs*, *7*(1). https://doi.org/10.3390/md7010024

Hoops, S., Gauges, R., Lee, C., Pahle, J., Simus, N., Singhal, M., Xu, L., Mendes, P., & Kummer, U. (2006). COPASI - A COmplex PAthway SImulator. *Bioinformatics*, *22*(24). https://doi.org/10.1093/bioinformatics/btl485

Huang, S.-H., Tseng, J.-C., Lin, C.-Y., Kuo, Y.-Y., Wang, B.-J., Kao, Y.-H., Muller, C. J. F., Joubert, E., & Chuu, C.-P. (2019). Rooibos suppresses proliferation of castration-resistant prostate cancer cells via inhibition of Akt signaling. *Phytomedicine*, *64*, 153068.

Iivanainen, E. K., Martikainen, P. J., Räisänen, M. L., & Katila, M. L. (1997). Mycobacteria in boreal coniferous forest soils. *FEMS Microbiology Ecology*, *23*(4). https://doi.org/10.1016/S0168-6496(97)00040-8

Iivanainen, E., Sallantaus, T., Katila, M.-L., & Martikainen, P. J. (1999). Mycobacteria in Runoff Waters from Natural and Drained Peatlands. *Journal of Environmental Quality*, *28*(4). https://doi.org/10.2134/jeq1999.00472425002800040025x

Inanami, O., Asanuma, T., Inukai, N., Jin, T., Shimokawa, S., Kasai, N., Nakano, M., Sato, F., & Kuwabara, M. (1995). The suppression of age-related accumulation of lipid peroxides in rat brain by administration of Rooibos tea (Aspalathus linearis). *Neuroscience Letters*, *196*(1–2), 85–88.

Ivanova, A. A., Beletsky, A. V., Rakitin, A. L., Kadnikov, V. V., Philippov, D. A., Mardanov, A. V., Ravin, N. V., & Dedysh, S. N. (2020). Closely located but totally distinct: Highly contrasting prokaryotic diversity patterns in raised bogs and eutrophic fens. *Microorganisms*, *8*(4), 484.

Jeon, S. J., & Park, J. H. (2020). Refolding, characterization, and dye decolorization ability of a highly thermostable laccase from Geobacillus sp. JS12. *Protein Expression and Purification*, *173*. https://doi.org/10.1016/j.pep.2020.105646

Job, N., & Ellery, W. N. (2013). Halting degradation of Southern Cape peatlands in agricultural landscapes. In *Policy Brief* (Issue 8).

Joubert, E., Gelderblom, W. C. A., & Beer, D. D. (2009). Phenolic contribution of South African herbal teas to a healthy diet. *Natural Product Communications*, *4*(5), 1934578X0900400507.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kazda, J. (2000). Mycobacteria as habitat microorganisms of sphagnum bog vegetation. In *The Ecology of Mycobacteria*. https://doi.org/10.1007/978-94-011-4102-4_5

Kennedy, A. C. (1999). Microbial diversity in agroecosystem quality. *Biodiversity in Agroecosystems*.

Kim, Y., Maltseva, N., Dementieva, I., Collart, F., Holzle, D., & Joachimiak, A. (2006). Crystal structure of hypothetical protein YfiH from *Shigella flexneri* at 2 Å resolution. *Proteins: Structure, Function, and Bioinformatics*, *63*(4), 1097–1101. https://doi.org/10.1002/prot.20589

Komori, H., & Higuchi, Y. (2015). Structural insights into the O_2 reduction mechanism of multicopper oxidase. *Journal of Biochemistry*, *158*(4), 293–298. https://doi.org/10.1093/jb/mvv079

Kumar, S. V. S., Phale, P. S., Durani, S., & Wangikar, P. P. (2003). Combined sequence and structure analysis of the fungal laccase family. *Biotechnology and Bioengineering*, *83*(4). https://doi.org/10.1002/bit.10681

Kusai, N. A., Ayob, Z., Maidin, M. S. T., Safari, S., & Ali, S. R. A. (2018). Characterization of fungi from different ecosystems of tropical peat in Sarawak, Malaysia. *Rendiconti Lincei*, *29*(2). https://doi.org/10.1007/s12210-018-0685-8

Lal, D., Verma, M., Behura, S. K., & Lal, R. (2016). Codon usage bias in phylum Actinobacteria: Relevance to environmental adaptation and host pathogenicity. *Research in Microbiology*, *167*(8), 669–677. https://doi.org/10.1016/j.resmic.2016.06.003

Laskowski, R. A., Hutchinson, E. G., Michie, A. D., Wallace, A. C., Jones, M. L., & Thornton, J. M. (1997). PDBsum: A web-based database of summaries and analyses of all PDB structures. *Trends in Biochemical Sciences*, *22*(12), 488–490. https://doi.org/10.1016/S0968-0004(97)01140-7

Lauterbach, S., Dienhart, H., Range, J., Malzacher, S., Spöring, J. D., Rother, D., Pinto, M. F., Martins, P., Lagerman, C. E., Bommarius, A. S., Høst, A. V., Woodley, J. M., Ngubane, S., Kudanga, T., Bergmann, F. T., Rohwer, J. M., Iglezakis, D., Weidemann, A., Wittig, U., … Pleiss, J. (2023). EnzymeML: seamless data flow and modeling of enzymatic data. *Nature Methods*, *20*(3). https://doi.org/10.1038/s41592-022-01763-1

Law, J. W. F., Letchumanan, V., Tan, L. T. H., Ser, H. L., Goh, B. H., & Lee, L. H. (2020). The Rising of "Modern Actinobacteria" Era. *Progress in Microbes and Molecular Biology*, *3*(1). https://doi.org/10.36877/pmmb.a0000064

Lee, L. H., Chan, K. G., Stach, J., Wellington, E. M. H., & Goh, B. H. (2018). Editorial: The search for biological active agent(s) from actinobacteria. *Frontiers in Microbiology*, *9*(MAY). https://doi.org/10.3389/fmicb.2018.00824

Leff, J. W., Jones, S. E., Prober, S. M., Barberán, A., Borer, E. T., Firn, J. L., Harpole, W. S., Hobbie, S. E., Hofmockel, K. S., Knops, J. M. H., McCulley, R. L., Pierre, K. L., Risch, A. C., Seabloom, E. W., Schütz, M., Steenbock, C., Stevens, C. J., & Fierer, N. (2015). Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(35). https://doi.org/10.1073/pnas.1508382112

Lerat, S., Forest, M., Lauzier, A., Grondin, G., Lacelle, S., & Beaulieu, C. (2012). Potato Suberin Induces Differentiation and Secondary Metabolism in the Genus *Streptomyces*. *Microbes and Environments*, *27*(1), 36–42. https://doi.org/10.1264/jsme2.ME11282

Lew, S., Glińska-Lewczuk, K., & Ziembińska-Buczyńska, A. (2018). Prokaryotic community composition affected by seasonal changes in physicochemical properties of water in peat bog lakes. *Water (Switzerland)*, *10*(4). https://doi.org/10.3390/w10040485

Li, Q., Chen, X., Jiang, Y., & Jiang, C. (2016). Morphological Identification of Actinobacteria. In D. Dhanasekaran & Y. Jiang (Eds.), *Actinobacteria—Basics and Biotechnological Applications*. InTech. https://doi.org/10.5772/61461

Li, Z., Kessler, W., van den Heuvel, J., & Rinas, U. (2011). Simple defined autoinduction medium for high-level recombinant protein production using T7-based Escherichia coli expression systems. *Applied Microbiology and Biotechnology*, *91*(4), 1203–1213. https://doi.org/10.1007/s00253-011-3407-z

Lischer, H. E. L., & Shimizu, K. K. (2017). Reference-guided de novo assembly approach improves genome reconstruction for related species. *BMC Bioinformatics*, *18*(1), 474. https://doi.org/10.1186/s12859-017-1911-6

Liu, B., Talukder, M. J. H., Terhonen, E., Lampela, M., Vasander, H., Sun, H., & Asiegbu, F. (2020). The microbial diversity and structure in peatland forest in Indonesia. *Soil Use and Management*, *36*(1). https://doi.org/10.1111/sum.12543

Liu, Y., Huang, J., & Zhang, X. (2009). Decolorization and biodegradation of remazol brilliant blue R by bilirubin oxidase. *Journal of Bioscience and Bioengineering*, *108*(6). https://doi.org/10.1016/j.jbiosc.2009.06.001

Locey, K. J., Fisk, M. C., & Lennon, J. T. (2017). Microscale insight into microbial seed banks. *Frontiers in Microbiology*, *7*(JAN). https://doi.org/10.3389/fmicb.2016.02040

Low, A. B., & Rebelo, A. G. (1998). *Vegetation of South Africa, Lesotho and Swaziland: A companion to the vegetation map of South Africa, Lesotho and Swaziland*. Pretoria (South Africa) Dept. of Environmental Affairs and Tourism.

Lu, L., Zeng, G., Fan, C., Ren, X., Wang, C., Zhao, Q., Zhang, J., Chen, M., Chen, A., & Jiang, M. (2013). Characterization of a laccase-like multicopper oxidase from newly isolated Streptomyces sp. C1 in agricultural waste compost and enzymatic decolorization of azo dyes. *Biochemical Engineering Journal*, *72*, 70–76. https://doi.org/10.1016/j.bej.2013.01.004

Lu, Y., Zhou, G., Ewald, J., Pang, Z., Shiri, T., & Xia, J. (2023). MicrobiomeAnalyst 2.0: Comprehensive statistical, functional and integrative analysis of microbiome data. *Nucleic Acids Research*, *51*(1 W). https://doi.org/10.1093/nar/gkad407

Lüdemann, H., & Conrad, R. (2000). Molecular Retrieval of Large 16S rRNA Gene Fragments from an Italian Rice Paddy Soil Affiliated with the Class Actinobacteria. *Systematic and Applied Microbiology*, *23*(4), 582–584. https://doi.org/10.1016/S0723-2020(00)80033-8

Ludwig, W., Euzéby, J., & Whitman, W. B. (2012). Taxonomic outline of the phylum Actinobacteria. In *Bergey's Manual® of Systematic Bacteriology*. https://doi.org/10.1007/978-0-387-68233-4_2

Macagnan, D., Romeiro, R. D. S., Souza, J. T. D., & Pomella, A. W. V. (2006). Isolation of actinomycetes and endospore-forming bacteria from the cacao pod surface and their antagonistic activity against the witches' broom and black pod pathogens. *Phytoparasitica*, *34*(2). https://doi.org/10.1007/BF02981312

Machczynski, M. C., Vijgenboom, E., Samyn, B., & Canters, G. W. (2004). Characterization of SLAC: A small laccase from Streptomyces coelicolor with unprecedented activity. *Protein Science*, *13*(9). https://doi.org/10.1110/ps.04759104

Magcwebeba, T. U., Swart, P., Swanevelder, S., Joubert, E., & Gelderblom, W. C. A. (2016). In vitro chemopreventive properties of green tea, rooibos and honeybush extracts in skin cells. *Molecules*, *21*(12), 1622.

Makhalanyane, T. P., Goethem, M. W. V., & Cowan, D. A. (2016). Microbial diversity and functional capacity in polar soils. *Current Opinion in Biotechnology*, *38*, 159–166. https://doi.org/10.1016/j.copbio.2016.01.011

Maldonado, L. A., Fenical, W., Jensen, P. R., Kauffman, C. A., Mincer, T. J., Ward, A. C., Bull, A. T., & Goodfellow, M. (2005). Salinispora arenicola gen. Nov., sp. Nov. And Salinispora tropica sp. Nov., obligate marine actinomycetes belonging to the family Micromonosporaceae. *International Journal of Systematic and Evolutionary Microbiology*, *55*(5). https://doi.org/10.1099/ijs.0.63625-0

Malzacher, S., Range, J., Halupczok, C., Pleiss, J., & Rother, D. (2020). BioCatHub, a graphical user interface for standardized data acquisition in biocatalysis. *Chemie Ingenieur Technik*, *92*(9). https://doi.org/10.1002/cite.202055297

Mandel, M., & Marmur, J. (1968). [109] Use of ultraviolet absorbance-temperature profile for determining the guanine plus cytosine content of DNA. In *Methods in Enzymology* (Vol. 12, pp. 195–206). Elsevier. https://doi.org/10.1016/0076-6879(67)12133-2

Manning, J. (2018). *Field guide to fynbos*. Penguin Random House South Africa.

Mano, N. (2012). Features and applications of bilirubin oxidases. In *Applied Microbiology and Biotechnology* (Vol. 96, Issue 2). https://doi.org/10.1007/s00253-012-4312-9

Marie-Vivien, D., & Biénabe, E. (2017). The Multifaceted Role of the State in the Protection of Geographical Indications: A Worldwide Review. *World Development*, *98*. https://doi.org/10.1016/j.worlddev.2017.04.035

Marnewick, J. L., Rautenbach, F., Venter, I., Neethling, H., Blackhurst, D. M., Wolmarans, P., & Macharia, M. (2011). Effects of rooibos (Aspalathus linearis) on oxidative stress and biochemical parameters in adults at risk for cardiovascular disease. *Journal of Ethnopharmacology*, *133*(1), 46–52.

Marnewick, J. L., Westhuizen, F. H. V. der, Joubert, E., Swanevelder, S., Swart, P., & Gelderblom, W. C. A. (2009). Chemoprotective properties of rooibos (Aspalathus linearis), honeybush (Cyclopia intermedia) herbal and green and black (Camellia sinensis) teas against cancer promotion induced by fumonisin B1 in rat liver. *Food and Chemical Toxicology*, *47*(1), 220–229.

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PloS One*, *8*(4), e61217.

Meadows, M. E., & Sugden, J. M. (1993). The late Quaternary palaeoecology of a floristic kingdom: The southwestern Cape South Africa. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *101*(3–4), 271–281.

Messerschmidt, A., & Huber, R. (1990). The blue oxidases, ascorbate oxidase, laccase and ceruloplasmin modelling and structural relationships. *European Journal of Biochemistry*, *187*(2), 341–352.

Minné, D., Stromin, J., Docrat, T., Engel-Hills, P., & Marnewick, J. L. (2023). The effects of tea polyphenols on emotional homeostasis: Understanding dementia risk through stress, mood, attention & sleep. *Clinical Nutrition ESPEN*.

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: Making protein folding accessible to all. *Nature Methods*, *19*(6), 679–682. https://doi.org/10.1038/s41592-022-01488-1

Mohammadipanah, F., & Wink, J. (2016). Actinobacteria from arid and desert habitats: Diversity and biological activity. *Frontiers in Microbiology*, *6*(JAN). https://doi.org/10.3389/fmicb.2015.01541

Montero-Calasanz, M. del C., Yaramis, A., Rohde, M., Schumann, P., Klenk, H. P., & Meier-Kolthoff, J. P. (2022). Genotype–phenotype correlations within the Geodermatophilaceae. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.975365

Moreno, A. D., Ibarra, D., Eugenio, M. E., & Tomás-Pejó, E. (2020). Laccases as versatile enzymes: From industrial uses to novel applications. In *Journal of Chemical Technology and Biotechnology* (Vol. 95, Issue 3). https://doi.org/10.1002/jctb.6224

Morozova, O. V., Shumakovich, G. P., Shleev, S. V., & Yaropolov, Y. I. (2007). Laccase-mediator systems and their applications: A review. *Applied Biochemistry and Microbiology*, *43*(5). https://doi.org/10.1134/S0003683807050055

Mucina, L., & Rutherford, M. C. (2006). The vegetation of South Africa, Lesotho and Swaziland. In *Strelitzia* (Vol. 19).

Muofhe, M. L., & Dakora, F. D. (2000). Modification of rhizosphere pH by the symbiotic legume Aspalathus linearis growing in a sandy acidic soil. *Australian Journal of Plant Physiology*, *27*(12). https://doi.org/10.1071/pp99198

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Fonseca, G. A. B. D., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853–858.

Nakamura, K., Kawabata, T., Yura, K., & Go, N. (2003). Novel types of two-domain multi-copper oxidases: Possible missing links in the evolution. *FEBS Letters*, *553*(3), 239–244.

Narnoliya, L. K., Agarwal, N., Patel, S. N., & Singh, S. P. (2019). Kinetic characterization of laccase from Bacillus atrophaeus, and its potential in juice clarification in free and immobilized forms. *Journal of Microbiology*, *57*(10). https://doi.org/10.1007/s12275-019-9170-z

Olivier, B. G., Rohwer, J. M., & Hofmeyr, J. H. S. (2005). Modelling cellular systems with PySCeS. *Bioinformatics*, *21*(4). https://doi.org/10.1093/bioinformatics/bti046

Oren, A., & Garrity, G. M. (2021). Candidatus list no. 2. Lists of names of prokaryotic candidatus taxa. *International Journal of Systematic and Evolutionary Microbiology*, *71*(4). https://doi.org/10.1099/ijsem.0.004671

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S. W., McGuire, A. D., Piao, S., Rautiainen, A., Sitch, S., & Hayes, D. (2011). A large and persistent carbon sink in the world's forests. *Science*, *333*(6045). https://doi.org/10.1126/science.1201609

Pankratov, T. A., Dedysh, S. N., & Zavarzin, G. A. (2006). The leading role of actinobacteria in aerobic cellulose degradation in Sphagnum peat bogs. *Doklady Biological Sciences*, *410*(1). https://doi.org/10.1134/S0012496606050243

Pantsi, W. G., Marnewick, J. L., Esterhuyse, A. J., Rautenbach, F., & Rooyen, J. V. (2011). Rooibos (Aspalathus linearis) offers cardiac protection against ischaemia/reperfusion in the isolated perfused rat heart. *Phytomedicine*, *18*(14), 1220–1228.

Parish, F., Sirin, A., Charman, D., Joosten, H., Minayeva, T., Silvius, M., & Stringer, L. (2008). Assessment on Peatland, Biodiversity and Climate Change. In *Global Environment Centre, Kuala Lumpur & Wetlands International, Wageningen* (Vol. 1).

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, *25*(7), 1043–1055. https://doi.org/10.1101/gr.186072.114

Parte, A. C., Sardà Carbasse, J., Meier-Kolthoff, J. P., Reimer, L. C., & Göker, M. (2020). List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. *International Journal of Systematic and Evolutionary Microbiology*, *70*(11), 5607–5612. https://doi.org/10.1099/ijsem.0.004332

Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B. L., Salazar, G. A., Bileschi, M. L., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D. H., Letunić, I., Marchler-Bauer, A., Mi, H., Natale, D. A., Orengo, C. A., Pandurangan, A. P., Rivoire, C., … Bateman, A. (2023). InterPro in 2022. *Nucleic Acids Research*, *51*(D1), D418–D427. https://doi.org/10.1093/nar/gkac993

Peters, S., Koschinsky, S., Schwieger, F., & Tebbe, C. C. (2000). Succession of Microbial Communities during Hot Composting as Detected by PCR–Single-Strand-Conformation Polymorphism-Based Genetic Profiles of Small-Subunit rRNA Genes. *Applied and Environmental Microbiology*, *66*(3), 930–936. https://doi.org/10.1128/AEM.66.3.930-936.2000

Pinto, M. F., Baici, A., Pereira, P. J. B., Macedo-Ribeiro, S., Pastore, A., Rocha, F., & Martins, P. M. (2021). interferENZY: A Web-Based Tool for Enzymatic Assay Validation and Standardized Kinetic Analysis. *Journal of Molecular Biology*, *433*(11). https://doi.org/10.1016/j.jmb.2020.07.025

Pop, M., Phillippy, A., Delcher, A. L., & Salzberg, S. L. (2004). Comparative genome assembly. *Briefings in Bioinformatics*, *5*(3), 237–248.

Postma, A. (2016). *Soil microbial communities associated with two commercially important plant species indigenous to the fynbos region of South Africa: Cyclopia spp.(honeybush) and Aspalathus linearis (rooibos).* Doctoral thesis, University of Stellenbosch, Stellenbosch, South Africa

Pretorius, L., & Smith, C. (2022). Aspalathus linearis (Rooibos) and agmatine may act synergistically to beneficially modulate intestinal tight junction integrity and inflammatory profile. *Pharmaceuticals*, *15*(9), 1097.

Prins, A. (2015). *The degradation of the endocrine disrupting chemical, bisphenol-A: a comparative study between fungal and bacterial laccases*. Masters thesis, University of the Western Cape, Cape Town, South Africa

Prins, A., Kleinsmidt, L., Khan, N., Kirby, B., Kudanga, T., Vollmer, J., Pleiss, J., Burton, S., & Roes-Hill, M. L. (2015). The effect of mutations near the T1 copper site on the biochemical characteristics of the small laccase from Streptomyces coelicolor A3(2). *Enzyme and Microbial Technology*, *68*. https://doi.org/10.1016/j.enzmictec.2014.10.003

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., & Korobeynikov, A. (2020). Using SPAdes De Novo Assembler. *Current Protocols in Bioinformatics*, *70*(1), e102. https://doi.org/10.1002/cpbi.102

Proemse, B. C., Eberhard, R. S., Sharples, C., Bowman, J. P., Richards, K., Comfort, M., & Barmuta, L. A. (2017). Stromatolites on the rise in peat-bound karstic wetlands. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-15507-1

Qin, S., Li, W. J., Dastager, S. G., & Hozzein, W. N. (2016). Editorial: Actinobacteria in special and extreme habitats: Diversity, function roles, and environmental adaptations. In *Frontiers in Microbiology* (Vol. 7, Issue SEP). https://doi.org/10.3389/fmicb.2016.01415

Range, J., Halupczok, C., Lohmann, J., Swainston, N., Kettner, C., Bergmann, F. T., Weidemann, A., Wittig, U., Schnell, S., & Pleiss, J. (2022). EnzymeML—a data exchange format for biocatalysis and enzymology. *FEBS Journal*, *289*(19). https://doi.org/10.1111/febs.16318

Rebelo, A. J., Morris, C., Meire, P., & Esler, K. J. (2019). Ecosystem services provided by South African palmiet wetlands: A case for investment in strategic water source areas. *Ecological Indicators*, *101*, 71–80.

Reiss, R., Ihssen, J., Richter, M., Eichhorn, E., Schilling, B., & Thöny-Meyer, L. (2013). Laccase versus Laccase-Like Multi-Copper Oxidase: A Comparative Study of Similar Enzymes with Diverse Substrate Spectra. *PLoS ONE*, *8*(6). https://doi.org/10.1371/journal.pone.0065633

Reiss, R., Ihssen, J., & Thöny-Meyer, L. (2011). Bacillus pumilus laccase: A heat stable enzyme with a wide substrate spectrum. *BMC Biotechnology*, *11*. https://doi.org/10.1186/1472-6750-11-9

Richards, M. B., Cowling, R. M., & Stock, W. D. (1995). Fynbos plant communities and vegetation–environment relationships in the Soetanysberg hills, Western Cape. *South African Journal of Botany*, *61*(6), 298–305. https://doi.org/10.1016/S0254-6299(15)30550-0

Richards, M. B., Cowling, R. M., & Stock, W. D. (1997). Soil nutrient dynamics and community boundaries in the fynbos vegetation of South Africa. *Plant Ecology*, *130*, 143–153.

Rodríguez-Couto, S. & Herrera, J. L. T. (2006). Industrial and biotechnological applications of laccases: A review. In *Biotechnology Advances* (Vol. 24, Issue 5). https://doi.org/10.1016/j.biotechadv.2006.04.003

Roes-Hill, M. L., & Meyers, P. R. (2009). Streptomyces polyantibioticus sp. Nov., isolated from the banks of a river. *International Journal of Systematic and Evolutionary Microbiology*, *59*(6). https://doi.org/10.1099/ijs.0.006171-0

Roes-Hill, M. L., & Prins, A. (2016). Biotechnological Potential of Oxidative Enzymes from Actinobacteria. In *Actinobacteria—Basics and Biotechnological Applications*. https://doi.org/10.5772/61321

Rossum, T. V., Kengen, S. W. M., & Oost, J. V. D. (2013). Reporter-based screening and selection of enzymes. *FEBS Journal*, *280*(13). https://doi.org/10.1111/febs.12281

Roth, S., & Spiess, A. C. (2015). Laccases for biorefinery applications: A critical review on challenges and perspectives. In *Bioprocess and Biosystems Engineering* (Vol. 38, Issue 12). https://doi.org/10.1007/s00449-015-1475-7

Roucher, A., Roussarie, E., Gauvin, R. M., Rouhana, J., Gounel, S., Stines-Chaumeil, C., Mano, N., & Backov, R. (2019). Bilirubin oxidase-based silica macrocellular robust catalyst for on line dyes degradation. *Enzyme and Microbial Technology*, *120*. https://doi.org/10.1016/j.enzmictec.2018.10.004

Roux, J. J. L., Keet, J. H., Mutiti, B., & Ellis, A. G. (2017). Cultivation may not dramatically alter rhizobial community diversity or structure associated with rooibos tea (Aspalathus linearis Burm.f.) in South Africa. *South African Journal of Botany*, *110*. https://doi.org/10.1016/j.sajb.2017.01.014

Rutherford, M. C., Mucina, L., & Powrie, L. W. (2006). Biomes and Bioregions of southern Africa. In *The vegetation of South Africa, Lesotho and Swaziland* (Vol. 19, Issue March 2015).

Sakurai, T., & Kataoka, K. (2007). Basic and applied features of multicopper oxidases, cueo, bilirubin oxidase, and laccase. *Chemical Record*, *7*(4). https://doi.org/10.1002/tcr.20125

Santhanam, N., Vivanco, J. M., Decker, S. R., & Reardon, K. F. (2011). Expression of industrially relevant laccases: Prokaryotic style. In *Trends in Biotechnology* (Vol. 29, Issue 10). https://doi.org/10.1016/j.tibtech.2011.04.005

Schäfer, J., Jäckel, U., & Kämpfer, P. (2010). Development of a new PCR primer system for selective amplification of Actinobacteria. *FEMS Microbiology Letters*, *311*(2). https://doi.org/10.1111/j.1574-6968.2010.02069.x

Schein, C. H., & Noteborn, M. H. M. (1988). Formation of Soluble Recombinant Proteins in Escherichia Coli is Favored by Lower Growth Temperature. *Nature Biotechnology*, *6*(3), 291–294. https://doi.org/10.1038/nbt0388-291

Schroeder, D., Chennells, R., Louw, C., Snyders, L., & Hodges, T. (2020). The Rooibos Benefit Sharing Agreement-Breaking New Ground with Respect, Honesty, Fairness, and Care. *Cambridge Quarterly of Healthcare Ethics*, *29*(2). https://doi.org/10.1017/S0963180119001075

Seong, C. N., Choi, J. H., & Baik, K. S. (2001). An Improved Selective Isolation of Rare Actinomycetes from Forest Soil. *Journal of Microbiology*, *39*(1).

Sharma, K. K., & Kuhad, R. C. (2009). An evidence of laccases in archaea. *Indian Journal of Microbiology*, *49*(2). https://doi.org/10.1007/s12088-009-0039-4

Sharma, V., Ayothiraman, S., & Dhakshinamoorthy, V. (2019). Production of highly thermo-tolerant laccase from novel thermophilic bacterium Bacillus sp. PC-3 and its application in functionalization of chitosan film. *Journal of Bioscience and Bioengineering*, *127*(6). https://doi.org/10.1016/j.jbiosc.2018.11.008

Shivlata, L., & Satyanarayana, T. (2015). Thermophilic and alkaliphilic Actinobacteria: Biology and potential applications. *Frontiers in Microbiology*, *6*(SEP). https://doi.org/10.3389/fmicb.2015.01014

Sigrist, C. J. A., Cerutti, L., Castro, E. D., Langendijk-Genevaux, P. S., Bulliard, V., Bairoch, A., & Hulo, N. (2009). PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Research*, *38*(SUPPL.1). https://doi.org/10.1093/nar/gkp885

Silberbauer, M. J., & King, J. M. (1991). Geographical trends in the water chemistry of wetlands in the south-western cape province, south africa. *Southern African Journal of Aquatic Sciences*, *17*(1–2). https://doi.org/10.1080/10183469.1991.9631315

Silva, L. J., Crevelin, E. J., Souza, D. T., Lacerda-Júnior, G. V., Oliveira, V. M. de, Ruiz, A. L. T. G., Rosa, L. H., Moraes, L. A. B., & Melo, I. S. (2020). Actinobacteria from Antarctica as a source for anticancer discovery. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-69786-2

Sirim, D., Wagner, F., Wang, L., Schmid, R. D., & Pleiss, J. (2011). The Laccase Engineering Database: A classification and analysis system for laccases and related multicopper oxidases. *Database*, *2011*. https://doi.org/10.1093/database/bar006

Skálová, T., Dohnálek, J., Østergaard, L. H., Østergaard, P. R., Kolenko, P., Dušková, J., Štěpánková, A., & Hašek, J. (2009). The Structure of the Small Laccase from Streptomyces coelicolor Reveals a Link between Laccases and Nitrite Reductases. *Journal of Molecular Biology*, *385*(4). https://doi.org/10.1016/j.jmb.2008.11.024

Skálová, T., Dušková, J., Hašek, J., Štěpánková, A., Kova, T., Østergaard, L. H., & Dohnálek, J. (2011). Structure of laccase from Streptomyces coelicolor after soaking with potassium hexacyanoferrate and at an improved resolution of 2.3 Å. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, *67*(1). https://doi.org/10.1107/S1744309110046099

Skene, K. R. (1998). Cluster roots: Some ecological considerations. In *Journal of Ecology* (Vol. 86, Issue 6). https://doi.org/10.1046/j.1365-2745.1998.00326.x

Smith, J. F. N., Botha, A., & Hardie, A. G. (2018). Role of soil quality in declining rooibos (Aspalathus linearis) tea yields in the Clanwilliam area, South Africa. *Soil Research*, *56*(3). https://doi.org/10.1071/SR17029

Solano, F., Lucas-Elío, P., López-Serrano, D., Fernandez, E., & Sanchez-Amat, A. (2001). Dimethoxyphenol oxidase activity of different microbial blue multicopper proteins. *FEMS Microbiology Letters*, *204*(1), 175–181. https://doi.org/10.1111/j.1574-6968.2001.tb10882.x

Solomon, E. I., Sundaram, U. M., & Machonkin, T. E. (1996). Multicopper oxidases and oxygenases. *Chemical Reviews*, *96*(7). https://doi.org/10.1021/cr9500460

Srebotnik, E., Messner, K., & Foisner, R. (1988). Penetrability of white rot-degraded pine wood by the lignin peroxidase of Phanerochaete chrysosporium. *Applied and Environmental Microbiology*, *54*(11). https://doi.org/10.1128/aem.54.11.2608-2614.1988

Stach, J. E. M., Bathe, S., Clapp, J. P., & Burns, R. G. (2001). PCR-SSCP comparison of 16S rDNA sequence diversity in soil DNA obtained using different isolation and purification methods. *FEMS Microbiology Ecology*, *36*(2–3), 139–151. https://doi.org/10.1111/j.1574-6941.2001.tb00834.x

Stach, J. E. M., Maldonado, L. A., Ward, A. C., Goodfellow, M., & Bull, A. T. (2003). New primers for the class *Actinobacteria*: Application to marine and terrestrial environments. *Environmental Microbiology*, *5*(10), 828–841. https://doi.org/10.1046/j.1462-2920.2003.00483.x

Stackebrandt, E., & Schumann, P. (2006). Introduction to the taxonomy of actinobacteria. *Prokaryotes*, *3*, 297–321. https://doi.org/10.1007/0-387-30743-5

Stafford, W. H. L., Baker, G. C., Brown, S. A., Burton, S. G., & Cowan, D. A. (2005). Bacterial diversity in the rhizosphere of Proteaceae species. *Environmental Microbiology*, *7*(11). https://doi.org/10.1111/j.1462-2920.2005.00929.x

Standley, L., Winterton, P., Marnewick, J. L., Gelderblom, W. C. A., Joubert, E., & Britz, T. J. (2001). Influence of processing stages on antimutagenic and antioxidant potentials of rooibos tea. *Journal of Agricultural and Food Chemistry*, *49*(1). https://doi.org/10.1021/jf000802d

Stewart, G. R., Pate, J. S., & Unkovich, M. (1993). Characteristics of inorganic nitrogen assimilation of plants in fire-prone Mediterranean-type vegetation. *Plant, Cell & Environment*, *16*(4), 351–363.

Stock, W. D., & Midgley, G. F. (1995). Ecosystem response to elevated CO2: Nutrient availability and nutrient cycling. In *Global change and mediterranean-type ecosystems* (pp. 326–342). Springer.

Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*, *41*(1), 207–234. https://doi.org/10.1016/j.pep.2005.01.016

Su, Y. S., Wu, M. D., Chen, J. J., Cheng, M. J., Kuo, Y. H., Chai, C. Y., & Kwan, A. L. (2021). Secondary metabolites with anti-inflammatory activities from one actinobacteria amycolatopsis Taiwanensis. *Molecules*, *26*(19). https://doi.org/10.3390/molecules26195765

Sun, H., Terhonen, E., Koskinen, K., Paulin, L., Kasanen, R., & Asiegbu, F. O. (2014). Bacterial diversity and community structure along different peat soils in boreal forest. *Applied Soil Ecology*, *74*. https://doi.org/10.1016/j.apsoil.2013.09.010

Suriya, J., Bharathiraja, S., Manivasagan, P., & Kim, S. K. (2016). Enzymes From Rare Actinobacterial Strains. In *Advances in Food and Nutrition Research* (Vol. 79). https://doi.org/10.1016/bs.afnr.2016.08.002

Teufel, F., Almagro Armenteros, J. J., Johansen, A. R., Gíslason, M. H., Pihl, S. I., Tsirigos, K. D., Winther, O., Brunak, S., Von Heijne, G., & Nielsen, H. (2022). SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology*, *40*(7), 1023–1025. https://doi.org/10.1038/s41587-021-01156-3

Tfaily, M. M., Hodgkins, S., Podgorski, D. C., Chanton, J. P., & Cooper, W. T. (2012). Comparison of dialysis and solid-phase extraction for isolation and concentration of dissolved organic matter prior to Fourier transform ion cyclotron resonance mass spectrometry. *Analytical and Bioanalytical Chemistry*, *404*(2), 447–457. https://doi.org/10.1007/s00216-012-6120-6

Thomas, S. A., & Brochu, F. (2022). Curation at the point of measurement and traceability of measurement workflows. *Measurement: Sensors*, *23*, 100399. https://doi.org/10.1016/j.measen.2022.100399

Thwaites, R. N., & Cowling, R. M. (1988). Soil-vegetation relationships on the Agulhas plain, South Africa. *Catena*, *15*(3–4), 333–345.

Tipton, K. F., Armstrong, R. N., Bakker, B. M., Bairoch, A., Cornish-Bowden, A., Halling, P. J., Hofmeyr, J.-H., Leyh, T. S., Kettner, C., Raushel, F. M., Rohwer, J., Schomburg, D., & Steinbeck, C. (2014). Standards for

Reporting Enzyme Data: The STRENDA Consortium: What it aims to do and why it should be helpful. *Perspectives in Science*, *1*(1–6). https://doi.org/10.1016/j.pisc.2014.02.012

Torres, E., Bustos-Jaimes, I., & Borgne, S. L. (2003). Potential use of oxidative enzymes for the detoxification of organic pollutants. *Applied Catalysis B: Environmental*, *46*(1). https://doi.org/10.1016/S0926-3373(03)00228-5

Troskie, D., & Biénabe, E. (2013). Institution building and local industry dynamics: Lessons from the rooibos GI initiative. In *Developing Geographical Indications in the South: The Southern African Experience*. https://doi.org/10.1007/978-94-007-6748-5_5

Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G. F., Chater, K. F., & Van Sinderen, D. (2007). Genomics of *Actinobacteria*: Tracing the Evolutionary History of an Ancient Phylum. *Microbiology and Molecular Biology Reviews*, *71*(3), 495–548. https://doi.org/10.1128/MMBR.00005-07

Vollmers, J., Wiegand, S., & Kaster, A.-K. (2017). Comparing and Evaluating Metagenome Assembly Tools from a Microbiologist's Perspective—Not Only Size Matters! *PLOS ONE*, *12*(1), e0169662. https://doi.org/10.1371/journal.pone.0169662

Weels, S. S. L., Welz, P. J., Prins, A., & Roes-Hill, M. L. (2022). Impact of Physicochemical Parameters on the Diversity and Distribution of Microbial Communities Associated with Three South African Peatlands. *Microorganisms*, *10*(11). https://doi.org/10.3390/microorganisms10112103

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., Santos, L. B. da S., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). Comment: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*. https://doi.org/10.1038/sdata.2016.18

Wittig, U., Kania, R., Golebiewski, M., Rey, M., Shi, L., Jong, L., Algaa, E., Weidemann, A., Sauer-Danzwith, H., Mir, S., Krebs, O., Bittkowski, M., Wetsch, E., Rojas, I., & Müller, W. (2012). SABIO-RK - Database for biochemical reaction kinetics. *Nucleic Acids Research*, *40*(D1). https://doi.org/10.1093/nar/gkr1046

Wittig, U., Rey, M., Weidemann, A., Kania, R., & Müller, W. (2018). SABIO-RK: An updated resource for manually curated biochemical reaction kinetics. *Nucleic Acids Research*, *46*(D1). https://doi.org/10.1093/nar/gkx1065

Wynberg, R. (2023). Biopiracy: Crying wolf or a lever for equity and conservation? *Research Policy*, *52*(2). https://doi.org/10.1016/j.respol.2022.104674

Yang, J., Liu, D., Wang, X., Ji, C., Cheng, F., Liu, B., Hu, Z., Chen, S., Pental, D., Ju, Y., Yao, P., Li, X., Xie, K., Zhang, J., Wang, J., Liu, F., Ma, W., Shopan, J., Zheng, H., … Zhang, M. (2016). The genome sequence of allopolyploid Brassica juncea and analysis of differential homoeolog gene expression influencing selection. *Nature Genetics*, *48*(10), 1225–1232. https://doi.org/10.1038/ng.3657

Yao, K., Wang, F. Q., Zhang, H. C., & Wei, D. Z. (2013). Identification and engineering of cholesterol oxidases involved in the initial step of sterols catabolism in Mycobacterium neoaurum. *Metabolic Engineering*, *15*(1). https://doi.org/10.1016/j.ymben.2012.10.005

Yoon, S. H., Ha, S. M., Kwon, S., Lim, J., Kim, Y., Seo, H., & Chun, J. (2017). Introducing EzBioCloud: A taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *International Journal of Systematic and Evolutionary Microbiology*, *67*(5). https://doi.org/10.1099/ijsem.0.001755

Yu, Z., Loisel, J., Brosseau, D. P., Beilman, D. W., & Hunt, S. J. (2010). Global peatland dynamics since the Last Glacial Maximum. *Geophysical Research Letters*, *37*(13). https://doi.org/10.1029/2010GL043584

Zhang, J., & Zhang, L. (2011). Improvement of an isolation medium for actinomycetes. *Modern Applied Science*, *5*(2). https://doi.org/10.5539/mas.v5n2p124

Zhang, X., Liu, Y., Yan, K., & Wu, H. (2007). Decolorization of anthraquinone-type dye by bilirubin oxidase-producing nonligninolytic fungus Myrothecium sp. IMER1. *Journal of Bioscience and Bioengineering*, *104*(2). https://doi.org/10.1263/jbb.104.104

Zhang, Z., Wang, J., Wang, J., Wang, J., & Li, Y. (2020). Estimate of the sequenced proportion of the global prokaryotic genome. *Microbiome*, *8*(1). https://doi.org/10.1186/s40168-020-00903-z

APPENDIX A

Table A1: Sequencing barcodes and linker primer sequences (custom Com2x-Ac1186r bTEFAP® - MrDNA, Shallowater, Texas, USA)

Sample	Barcode Sequence	Linker Primer Sequence	Source	Site	Sample Date
1S.S1P1	CGTAAGGC	AAACTCAAAGGAATTGACGG	Soil	Clanwilliam	August 2016
1S.S1P2	CGTAATTG	AAACTCAAAGGAATTGACGG	Soil	Clanwilliam	August 2016
1S.S2P1	CGTACTAG	AAACTCAAAGGAATTGACGG	Soil	Clanwilliam	August 2016
1S.S2P2	CGTAGCTT	AAACTCAAAGGAATTGACGG	Soil	Clanwilliam	August 2016
1S.S3P1	CGTAGTGT	AAACTCAAAGGAATTGACGG	Soil	Clanwilliam	August 2016
1S.S3P2	CGTATCCA	AAACTCAAAGGAATTGACGG	Soil	Clanwilliam	August 2016
2S.S1P1	CGTATGTG	AAACTCAAAGGAATTGACGG	Soil	Clanwilliam	August 2016
2S.S1P2	CGTCAAGC	AAACTCAAAGGAATTGACGG	Soil	Clanwilliam	August 2016
2S.S2P1	CGTCATGC	AAACTCAAAGGAATTGACGG	Soil	Clanwilliam	August 2016
2S.S2P2	CGTCCCAG	AAACTCAAAGGAATTGACGG	Soil	Clanwilliam	August 2016
2S.S3P1	CGTCGAAT	AAACTCAAAGGAATTGACGG	Soil	Clanwilliam	August 2016
2S.S3P2	CGTCGGAG	AAACTCAAAGGAATTGACGG	Soil	Clanwilliam	August 2016
W.1A1	CGTAAGGC	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.1A2	CGTAATTG	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.1A3	CGTACTAG	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.1B1	CGTAGCTT	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.1B2	CGTAGTGT	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.1B3	CGTATCCA	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.1C1	CGTATGTG	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.1C2	CGTCAAGC	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.1C3	CGTCATGC	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.2A1	CGTCCCAG	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.2A2	CGTCGAAT	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.2A3	CGTCGGAG	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.2B1	CGTCGTGC	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.2B2	CGTCTCAT	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.2B3	CGTCTGTC	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.2C1	CGTGAAGT	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.2C2	CGTGATTG	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018
W.2C3	CGTGCCGC	AAACTCAAAGGAATTGACGG	Soil	Springfield	September 2018

Table A2: Good's coverage for rarefied library (Rooibos samples)

sample	no.singleton	no.seqs	goods
1S.S1P2	167	62750	99.7338645418327
1S.S2P1	137	62750	99.7816733067729
1S.S2P2	135	62750	99.7848605577689
1S.S3P1	175	62750	99.7211155378486
1S.S3P2	111	62750	99.8231075697211
2S.S1P1	173	62750	99.7243027888446
2S.S1P2	119	62750	99.8103585657371
2S.S2P1	139	62750	99.7784860557769
2S.S2P2	126	62750	99.799203187251
2S.S3P1	141	62750	99.7752988047809
2S.S3P2	184	62750	99.7067729083665

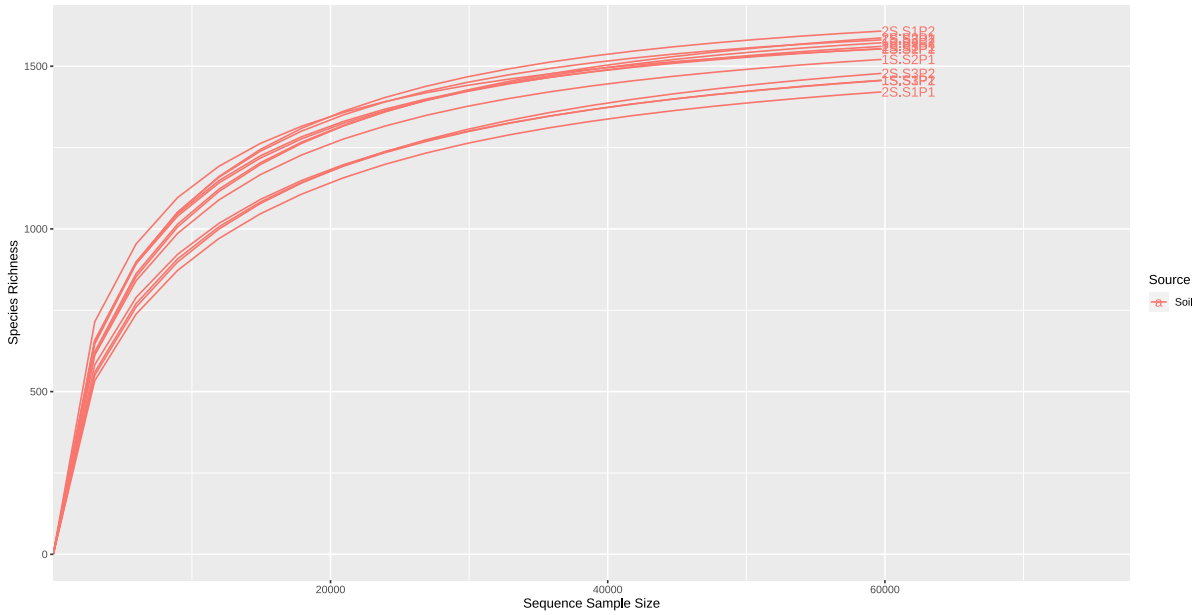


Figure A1: Rarefaction curve representing sequencing coverage for soil samples collected from the Rooibos environments.

Table A3: Good's coverage for rarefied library (Springfield samples)

sample	no.singleton	no.seqs	goods
W.1A1	116	23665	99.5098246355377
W.1A2	119	23665	99.4971476864568
W.1A3	94	23665	99.6027889287978
W.1B1	92	23665	99.6112402281851
W.1B2	95	23665	99.5985632791042
W.1B3	85	23665	99.6408197760406
W.1C1	88	23665	99.6281428269596
W.1C2	100	23665	99.577435030636
W.1C3	86	23665	99.6365941263469
W.2A1	96	23665	99.5943376294105
W.2A2	87	23665	99.6323684766533
W.2A3	117	23665	99.5055989858441
W.2B1	121	23665	99.4886963870695
W.2B2	110	23665	99.5351785336996
W.2B3	131	23665	99.4464398901331
W.2C1	121	23665	99.4886963870695
W.2C2	154	23665	99.3492499471794
W.2C3	132	23665	99.4422142404395

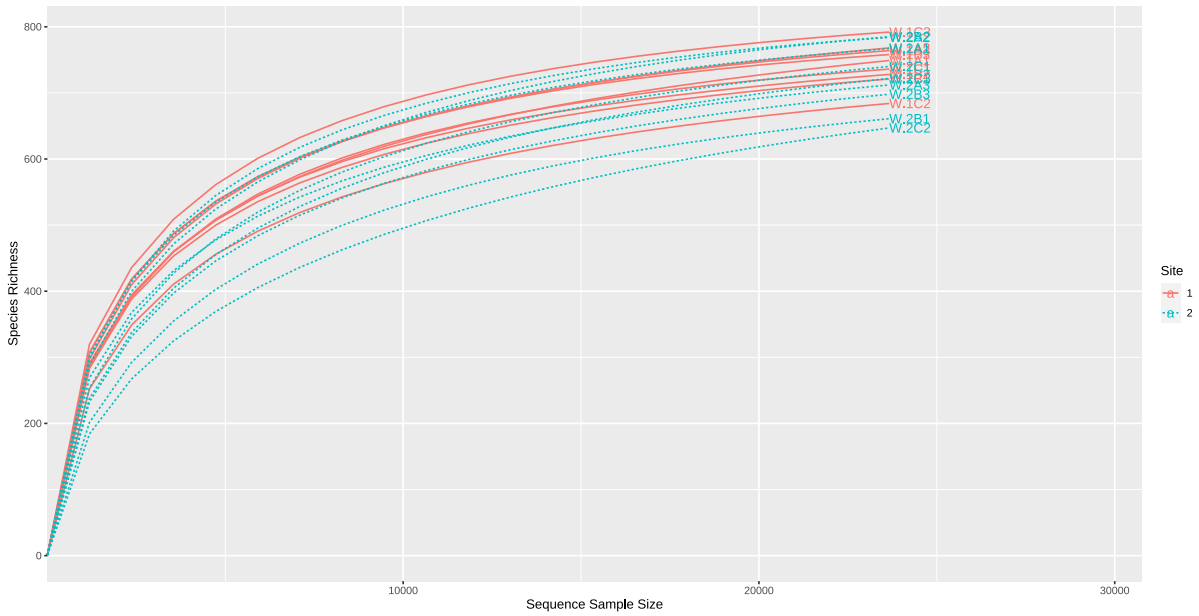


Figure A2: Rarefaction curve representing sequencing coverage for soil samples from the Springfield environment.

Table A4: Relative abundances of actinobacteria detected in Rooibos samples

	1S.S1P1	2S.S1P1	1S.S2P2	1S.S1P2	1S.S3P2	1S.S3P1	2S.S1P2	1S.S2P1	2S.S3P1	2S.S3P2	2S.S2P2	2S.S2P1
Source	Soil	Soil	Soil	Soil	Soil	Soil	Soil	Soil	Soil	Soil	Soil	Soil
sample	1S.S1P1	2S.S1P1	1S.S2P2	1S.S1P2	1S.S3P2	1S.S3P1	2S.S1P2	1S.S2P1	2S.S3P1	2S.S3P2	2S.S2P2	2S.S2P1
Mycobacteriaceae	0.14790139	0.21190864	0.09673251	0.19747596	0.10161612	0.13796633	0.21348616	0.27771158	0.15780288	0.28580059	0.20778022	0.1482706
Pseudonocardiaceae	0.1179955	0.2351855	0.1973081	0.2298488	0.188917	0.2921275	0.2314599	0.1039488	0.1743837	0.1938174	0.1806266	0.233071
Frankiaceae	0.10631514	0.03267491	0.05051437	0.06746438	0.06115428	0.08124255	0.05955997	0.08678067	0.05499522	0.04469096	0.09403058	0.08166211
Geodermatophilaceae	0.10510682	0.0571769	0.13596926	0.03233927	0.16847635	0.06437646	0.13006193	0.17985467	0.21756423	0.1031433	0.16134392	0.1859466
AF498716_f	0.10084414	0.26020776	0.04697333	0.25896588	0.05160522	0.12705792	0.09262087	0.06936077	0.03932066	0.18740665	0.05229329	0.04902076
Streptomycetaceae	0.0905231	0.06769933	0.26576267	0.07662745	0.17626328	0.07449611	0.08136003	0.11207142	0.06994814	0.06420864	0.09243627	0.06736369
Kineosporiaceae	0.079094433	0.011344756	0.058922248	0.006310101	0.079832849	0.017856244	0.042391797	0.016849313	0.036853676	0.006008022	0.035242586	0.061724873
Microbacteriaceae	0.042307886	0.012804806	0.028462584	0.005085002	0.028244416	0.020323225	0.026213771	0.037122191	0.061657744	0.035712488	0.021699364	0.024485207
Micromonosporaceae	0.040294024	0.005001091	0.014785104	0.005739507	0.017654858	0.016866095	0.013157232	0.006746438	0.04410358	0.002366288	0.013878866	0.008944904
Nocardioidaceae	0.036484468	0.008206488	0.024703375	0.009230201	0.030694615	0.023545404	0.0304261	0.029956198	0.034739121	0.007165993	0.044288184	0.031399466
Acidimicrobiaceae	0.027019316	0.010807727	0.014130599	0.015640996	0.017168174	0.026213771	0.009079162	0.006712874	0.009095944	0.017889808	0.012234212	0.008827429
Nakamurellaceae	0.019215601	0.001342575	0.00944837	0.001342575	0.013425747	0.001275446	0.003574605	0.006159062	0.007233121	0.001879605	0.013929213	0.033799319
Iamiaceae	0.013711044	0.006091933	0.003272526	0.010572776	0.003322872	0.003876685	0.003943813	0.002433417	0.002433417	0.002752278	0.003004011	0.002433417
Propionibacteriaceae	0.013643916	0.01444946	0.014415896	0.01513753	0.019165254	0.017805897	0.018846393	0.007719805	0.034336349	0.010287479	0.021078423	0.020843473
Sporichthyaceae	0.010388172	0.004245893	0.005017873	0.018359709	0.005806636	0.025173276	0.008659607	0.008105795	0.005269606	0.005856982	0.008995251	0.007048517
Actinospicaceae	0.006024804	0.039924816	0.002131337	0.02005471	0.002718714	0.026885059	0.005252824	0.003423566	0.006780002	0.007199557	0.004363368	0.005118566
Micrococcaceae	0.004984309	0.002181684	0.006947824	0.002685149	0.00652827	0.00599124	0.005051437	0.003238962	0.010723816	0.002970447	0.003541041	0.002567674
Microthrixaceae	0.004665447	0.000369208	0.000671287	0.000822327	0.000604159	0.000503466	0.001208317	0.001023713	0.000352426	0.000167822	0.000570594	0.000503466
Thermomonosporaceae	0.004262675	0.003390001	0.002517328	0.003339655	0.003087922	0.007568765	0.003255744	0.025878128	0.002903318	0.002685149	0.004547972	0.002450199
FJ478790_f	0.004078071	0.00253411	0.002349506	0.007602329	0.002265595	0.002651585	0.00092302	0.000855891	0.001879605	0.000520248	0.001661436	0.00107406
HQ910322_f	0.003708863	0.000788763	0.002802625	0.000486683	0.003859902	0.00360817	0.001057278	0.003792774	0.002802625	0.000755198	0.005554903	0.006024804
Nocardiaceae	0.00330609	0.002013862	0.003591387	0.003272526	0.00384312	0.002097773	0.00345713	0.001241882	0.003910249	0.002013862	0.002785843	0.002366288
Cellulomonadaceae	0.003289308	0.000285297	0.001275446	0.000268515	0.001124406	0.000402772	0.000553812	0.000704852	0.001157971	0.000973367	0.001426486	0.001023713
Ilumatobacter_f	0.003121486	0.000268515	0.000721634	0.000369208	0.001040495	0.000117475	0.000553812	0.000436337	0.000436337	0.000184604	0.000302079	0.000302079

GQ088405_f	0.00199708	0.000637723	0.000771981	0.000469901	0.000906238	0.001392921	0.001275446	0.000973367	0.000704852	0.000704852	0.001124406	0.001208317
Intrasporangiaceae	0.001527179	0.000503466	0.001090842	0.000503466	0.001376139	0.002785843	0.000973367	0.000503466	0.001896387	0.000453119	0.000889456	0.000822327
Promicromonosporaceae	0.001409704	0.000486683	0.000906238	0.000604159	0.001090842	0.000771981	0.001241882	0.000889456	0.002634803	0.004329804	0.001426486	0.001208317
Sanguibacteraceae	0.001191535	0.000184604	0.000486683	0.000100693	0.000553812	0.000352426	0.000704852	0.000234951	0.00438015	0.000251733	0.000335644	0.000570594
Motilibacteraceae	0.001006931	0.001057278	0.002265595	0.000587376	0.001560743	0.002265595	0.00214812	0.000839109	0.003759209	0.000268515	0.001711783	0.003054358
AY234742_f	0.000956585	0.000520248	0.001392921	0.002399852	0.001527179	0.001778912	0.00384312	0.001006931	0.000721634	0.00130901	0.001359357	0.001745347
Cryptosporangiaceae	9.23E-04	4.36E-04	3.86E-04	5.03E-05	6.38E-04	8.39E-05	2.18E-04	1.34E-04	1.68E-04	2.35E-04	4.03E-04	2.85E-04
Antricoccus_f	3.86E-04	1.68E-05	4.53E-04	1.34E-04	2.69E-04	4.36E-04	3.02E-04	6.04E-04	5.03E-05	1.34E-04	1.46E-03	1.12E-03
Bogoriellaceae	3.86E-04	8.39E-05	6.55E-04	1.68E-05	4.7E-04	1.01E-04	1.34E-04	4.03E-04	4.03E-04	6.04E-04	3.02E-04	4.03E-04
Catenulisporaceae	3.69E-04	1.21E-03	1.17E-04	8.06E-04	1.85E-04	4.38E-03	1.85E-04	3.36E-04	5.87E-04	8.39E-05	1.85E-04	1.17E-04
FJ478799_f	0.000302079	0.000352426	0.000704852	0.000889456	0.000872674	0.00161109	0.00038599	0.00053703	0.001225099	0.00038599	0.001543961	0.002097773
Jiangellaceae	3.02E-04	5.03E-05	3.36E-04	0E+00	5.87E-04	0E+00	3.36E-05	0E+00	1.01E-04	0E+00	0E+00	1.68E-05
Dermatophilaceae	2.85E-04	3.36E-05	3.36E-04	3.36E-05	3.36E-04	1.51E-04	2.01E-04	4.7E-04	5.54E-04	1.01E-04	4.53E-04	6.71E-05
EF016806_f	2.01E-04	6.71E-05	1.51E-04	1.68E-05	2.35E-04	0E+00	1.68E-05	3.36E-05	5.87E-04	1.68E-05	0E+00	1.17E-04
Dermabacteraceae	1.68E-04	1.85E-03	5.03E-05	8.73E-04	1.17E-04	6.04E-04	2.85E-04	1.34E-04	2.35E-04	3.86E-04	2.01E-04	2.85E-04
Nocardiopsaceae	0.000167822	0.001376139	0.000302079	0.002265595	0.00053703	0.000721634	0.001443268	0.00068807	0.00053703	0.002819407	0.000469901	0.000218168
Brevibacteriaceae	8.39E-05	3.36E-05	5.03E-05	1.68E-05	2.01E-04	1.21E-03	1.34E-04	1.68E-05	1.01E-04	1.36E-03	1.68E-05	1.68E-04
Streptosporangiaceae	3.36E-05	1.85E-04	3.36E-05	1.07E-03	5.03E-05	1.31E-03	2.35E-04	5.03E-05	6.04E-04	8.39E-05	4.87E-04	1.85E-04
Corynebacteriaceae	1.68E-05	1.68E-05	8.39E-05	1.17E-04	6.71E-05	1.68E-05	8.39E-05	1.68E-05	6.71E-05	1.68E-05	1.68E-05	3.36E-05

Table A5: Relative abundance of actinobacteria detected in Springfield samples

	W.1B1	W.1C3	W.2C1	W.1A2	W.2B2	W.1C1	W.2B3	W.2C2	W.1C2	W.2A2	W.2A3	W.2C3	W.1A3	W.2A1	W.2B1	W.1B3	W.1A1	W.1B2
sample	W.1B1	W.1C3	W.2C1	W.1A2	W.2B2	W.1C1	W.2B3	W.2C2	W.1C2	W.2A2	W.2A3	W.2C3	W.1A3	W.2A1	W.2B1	W.1B3	W.1A1	W.1B2
newnewnew	one	one	one	one	one	one	one	one	one	one	one	one	one	one	one	one	one	one
Not_Assigned	0.3199493	0.2749789	0.3134827	0.2532122	0.3299662	0.2072697	0.4979279	0.3254861	0.1777684	0.3268385	0.2112426	0.5297971	0.3025782	0.3427303	0.4734573	0.2809806	0.3138208	0.2116223
f__Mycobacteriaceae	0.15976331	0.19057481	0.05536771	0.05925613	0.09852071	0.25287405	0.06699107	0.02633136	0.23469992	0.09801352	0.04678783	0.04890111	0.06656805	0.07163199	0.02797971	0.32087912	0.20240913	0.35042265
f__koll13	0.11420118	0.08786982	0.03524937	0.09598478	0.04459003	0.04729501	0.03989856	0.01622992	0.04661877	0.08516484	0.01652578	0.05316991	0.2149197	0.07079459	0.03064243	0.07548605	0.0453508	0.06191885
f__Actinosynnemataceae	0.107607777	0.007607777	0.001986475	0.003212172	0.00397295	0.065004227	0.001986475	0.001056636	0.022062553	0.003592561	0.00295858	0.002197802	0.0034235	0.004860524	0.001775148	0.008284024	0.002704987	0.013144548
f__AKIW874	0.10350803	0.03677092	0.17349958	0.02633136	0.22527473	0.17540152	0.21513102	0.17671175	0.31952663	0.02641589	0.02295013	0.14733728	0.01986475	0.02730347	0.19699915	0.01952663	0.04674556	0.15760778
f__Pseudonocardiaceae	0.03888419	0.04010989	0.07840237	0.13060017	0.11267963	0.04551986	0.05646661	0.12658495	0.05731192	0.15701606	0.03021978	0.07920541	0.06052409	0.13338969	0.18013525	0.03123415	0.10409975	0.0397295
f__Micromonosporaceae	0.035925613	0.042392223	0.007480981	0.094885883	0.010777684	0.03956044	0.011496196	0.005536771	0.050633981	0.021724429	0.027218935	0.030092984	0.040912933	0.016018597	0.007988166	0.051606086	0.10443787	0.031318681
f__Streptomycetaceae	0.033558749	0.136686391	0.013736264	0.019949281	0.036179205	0.050507185	0.040278952	0.012933221	0.030008453	0.029247675	0.008833474	0.028698225	0.041504649	0.044336433	0.014919696	0.083093829	0.043744717	0.025401522
f__C111	0.0228656	0.03820795	0.04573119	0.09395604	0.04915469	0.02480981	0.01893491	0.01935757	0.01048183	0.03799662	0.08981403	0.01830093	0.05845309	0.04801352	0.01069315	0.02278107	0.02379544	0.02721893
f__Nocardioidaceae	0.013947591	0.031910397	0.046069315	0.056593407	0.025866441	0.021048183	0.008537616	0.008791209	0.014539307	0.074978867	0.237954353	0.015469146	0.034023669	0.074852071	0.011411665	0.020540997	0.031952663	0.017159763
f__Geodermatophilaceae	0.008833474	0.011031276	0.051563821	0.014158918	0.014412511	0.009340659	0.004945055	0.005029586	0.003296703	0.046914624	0.192645816	0.009256128	0.011876585	0.036179205	0.005494505	0.007523246	0.007819104	0.008748943
f__Intrasporangiaceae	0.008030431	0.037404903	0.011158073	0.063186813	0.008579882	0.015300085	0.004649197	0.003677092	0.005283178	0.023710904	0.047548605	0.006086221	0.087700761	0.013778529	0.004057481	0.028402367	0.035460693	0.013060017
f__Propionibacteriaceae	0.005283178	0.001141167	0.000126796	0.000591716	0.001267963	0.003212172	0.000169062	0.000676247	0.004057481	0.000549451	0.000464492	0.000507185	0.000676247	0.000338123	0.000422654	0.001394759	0.000887574	0.0034235
f__Nocardiaceae	0.00490279	0.006086221	0.003677092	0.005494505	0.005029586	0.006677937	0.003846154	0.003550296	0.0034235	0.005748098	0.00249366	0.00249366	0.004480135	0.005156382	0.00202874	0.0079459	0.004353339	0.008579882
f__Micrococcaceae	0.004099746	0.004945055	0.004860524	0.005536771	0.005029586	0.003338969	0.003592561	0.002113271	0.003169907	0.017497887	0.007819104	0.009044801	0.008114962	0.059467456	0.00498732	0.006001691	0.005114117	0.002535926

f__Nocardiopsaceae	0.002451395	0.004353339	0.131065089	0.010397295	0.003592561	0.004649197	0.003169907	0.243153001	0.00202874	0.003381234	0.002409129	0.002578191	0.002831784	0.003677092	0.008241758	0.004395604	0.004057481	0.002916314
f__Kineosporiaceae	0.002240068	0.001944421	0.000929839	0.005240913	0.000887574	0.001141167	0.000464492	0.000760778	0.000676247	0.001352494	0.001479291	0.000549451	0.003465765	0.001479291	0.000422654	0.002071006	0.002366864	0.000760778
f__Cellulomonadaceae	0.001732883	0.005367709	0.001267963	0.007819104	0.001479291	0.008918005	0.000760778	0.000633981	0.001141167	0.001901944	0.012087912	0.001183432	0.006086222	0.001563821	0.001056636	0.005325444	0.001521555	0.005579036
f__JdFBGBact	0.001732883	0.001521555	0.003127642	0.002747253	0.007861369	0.001098901	0.012130178	0.002451395	0.001267963	0.000676247	0.002071006	0.004142012	0.001014437	0.001732883	0.003677092	0.001141167	0.000929839	0.000760778
f__Thermomonosporaceae	0.001606086	0.013060017	0.008157227	0.010101437	0.003465765	0.002493366	0.001732883	0.012214708	0.004015216	0.007142857	0.004184277	0.003677092	0.004015216	0.005452224	0.001817413	0.005494505	0.002620456	0.003465765
f__Microbacteriaceae	0.001310228	0.003888419	0.002240068	0.005156382	0.001183432	0.001732883	0.000591716	0.000760778	0.001014437	0.003000845	0.004733728	0.000422654	0.005409975	0.004268808	0.000464492	0.003000845	0.002451395	0.001521555
f__ZA3409c	0.001310228	0.001606086	0.001014437	0.003846154	0.001648352	0.001014437	0.001310228	0.000422654	0.000929839	0.001014437	0.001098901	0.001479291	0.001183432	0.004437787	0.001267963	0.000549451	0.000929839	0.001437025
f__Promicromonosporaceae	0.001225697	0.007988166	0.000676247	0.008495351	0.001310228	0.001563821	0.000718512	0.001563821	0.000676247	0.000803043	0.001732883	0.000464492	0.009256129	0.000887574	0.000929839	0.005071851	0.004649197	0.002535926
f__Actinopolysporaceae	0.001056636	0.001648352	0.001859679	0.001648352	0.002197802	0.001310228	0.000929839	0.000676247	0.000803043	0.016356672	0.000803043	0.002113271	0.001098901	0.019526627	0.001183432	0.001817413	0.002620457	0.001183432
f__Iamiaceae	0.000929839	0.002282333	0.001732883	0.004395604	0.001479291	0.001690617	0.001225697	0.000718512	0.000972105	0.003127642	0.006001691	0.000718512	0.002409129	0.001606086	0.000845309	0.001859679	0.001648352	0.001859679
f__Actinomycetaceae	0.000845309	0.004437787	0.000422654	0.007396645	0.000380389	0.002155537	0.000464492	0.000126796	0.000845309	0.000718512	0.000549451	0.000295858	0.003000845	0.000549451	0.000169062	0.001225697	0.001352494	0.002071006
f__Microthrixaceae	0.000760778	0.003127642	0.003296703	0.008622147	0.001732883	0.002704987	0.000803043	0.000380389	0.001056636	0.002240068	0.009087067	0.000760778	0.003930685	0.004437787	0.000845309	0.001183432	0.001479291	0.002747253
f__Dietziaceae	5.49E-04	4.23E-05	0E+00	1.27E-04	0E+00	0E+00	4.23E-05	4.23E-05	0E+00	1.69E-04	0E+00	8.45E-05	8.45E-05	0E+00	0E+00	1.69E-04	4.23E-05	0E+00
f__Frankiaceae	5.07E-04	3.38E-04	1.23E-03	5.49E-04	3.8E-04	4.65E-04	5.07E-04	1.73E-03	2.11E-04	1.39E-03	2.96E-04	5.07E-04	1.27E-04	1.06E-03	5.71E-03	2.96E-04	2.11E-04	4.23E-05
f__Williamsiaceae	2.54E-04	5.07E-04	2.54E-04	1.27E-04	1.69E-04	7.19E-04	8.45E-05	4.23E-05	1.39E-03	8.45E-05	4.23E-05	0E+00	3.38E-04	2.11E-04	4.23E-05	4.65E-04	4.23E-05	7.61E-04
f__EB1017	8.45E-05	1.27E-04	1.27E-04	2.54E-04	1.69E-04	1.69E-04	0E+00	8.45E-05	4.23E-05	1.69E-04	3.21E-03	8.45E-05	4.23E-05	4.23E-05	1.27E-04	4.23E-05	3.8E-04	0E+00
f__Gordoniaceae	4.23E-05	0E+00	0E+00	8.45E-05	4.23E-04	4.23E-05	1.27E-04	0E+00	0E+00	8.03E-04	8.45E-05	2.96E-04	0E+00	8.45E-05	0E+00	0E+00	0E+00	0E+00
f__Sporichthyaceae	0E+00	4.23E-05	2.11E-04	4.23E-05	3.38E-04	4.23E-05	8.45E-05	1.69E-04	4.23E-05	1.69E-04	4.56E-03	4.23E-05	8.45E-05	1.27E-04	2.11E-04	0E+00	0E+00	0E+00

f__Streptosporangiaceae	0E+00	0E+00	0E+00	0E+00	0E+00	9.3E-04	0E+00	0E+00	0E+00	8.45E-05	8.45E-05	4.23E-05	0E+00	0E+00	0E+00	2.11E-04	0E+00	4.65E-04

Table A6: Actinobacteria isolated from Rooibos sediment samples

Strain	Isolation Media	Site	Plant	Source Material
1-20	JCM	1	1	Soil
1-38	ISP2	1	1	Soil
1-39	ISP2	1	1	Soil
1-40	ISP2	1	1	Soil
1-47	JCM	1	1	Soil
1-48	ISP2	1	1	Soil
1-50	ISP2	1	1	Soil
1-52	ISP2	1	1	Soil
1-54	Rooibos Extract	1	1	Soil
1-55	Rooibos Extract	1	1	Soil
1-59	ISP2	1	1	Soil
1-60	ISP2	1	1	Soil
1-61	ISP2	1	1	Soil
1-66	R2A	1	1	Soil
1-70	SCN	1	1	Soil
1-71	SCN	1	1	Soil
1-72	SCN	1	1	Soil
1-73	R2A	1	1	Soil
1-76	R2A	1	1	Soil
1-8	R2A	1	1	Soil
2-23	JCM	1	2	Soil
2-34	JCM	1	2	Soil
2-36	JCM	1	2	Soil
2-37	ISP2	1	2	Soil
2-38	ISP2	1	2	Soil
2-39	ISP2	1	2	Soil
2-8	ISP2	1	2	Soil
3-12	CZ	2	1	Soil
3-30	SCN	2	1	Soil
3-33	R2A	2	1	Soil
3-34	R2A	2	1	Soil
3-36	R2A	2	1	Soil
3-48	ISP2	2	1	Soil
3-49	ISP2	2	1	Soil
3-50	ISP2	2	1	Soil
3-51	JCM	2	1	Soil
3-52	JCM	2	1	Soil
3-53	JCM	2	1	Soil
3-57	CZ	2	1	Soil
3-6	R2A	2	1	Soil

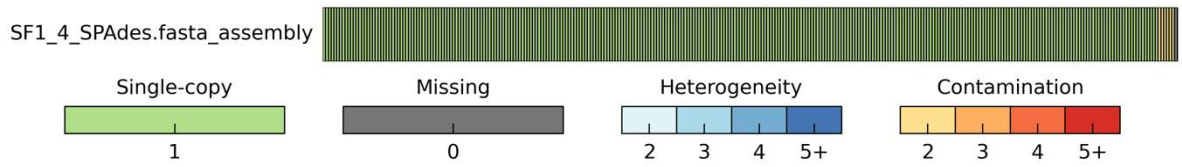
3-60	JCM	2	1	Soil
3-62	JCM	2	1	Soil
3-65	JCM	2	1	Soil
3-66	R2A	2	1	Soil
3-68	SCN	2	1	Soil
3-69	SCN	2	1	Soil
3-8	ISP2	2	1	Soil
4-11	JCM	2	2	Soil
4-13	JCM	2	2	Soil
4-19	JCM	2	2	Soil
4-22	ISP2	2	2	Soil
4-23	ISP2	2	2	Soil
4-25	ISP2	2	2	Soil
4-26	ISP2	2	2	Soil
4-27	ISP2	2	2	Soil
4-30	JCM	2	2	Soil
4-32	JCM	2	2	Soil
4-37	R2A	2	2	Soil
4-48	ISP2	2	2	Soil
4-50	ISP2	2	2	Soil
4-53	JCM	2	2	Soil
4-54	JCM	2	2	Soil
4-55	ISP2	2	2	Soil
4-56	ISP2	2	2	Soil
4-57	ISP2	2	2	Soil
4-58	ISP2	2	2	Soil
4-59	R2A	2	2	Soil
4-63	R2A	2	2	Soil
4-65	R2A	2	2	Soil
4-77	Rooibos Extract	2	2	Soil
4-8	JCM	2	2	Soil
4-9	JCM	2	2	Soil
5-12	R2A	3	1	Soil
5-14	R2A	3	1	Soil
5-15	R2A	3	1	Soil
5-2	JCM	3	1	Soil
5-3	JCM	3	1	Soil
5-31	ISP2	3	1	Soil
5-32	ISP2	3	1	Soil
5-33	ISP2	3	1	Soil
5-4	JCM	3	1	Soil
5-40	ISP2	3	1	Soil
5-42	ISP2	3	1	Soil
5-43	ISP2	3	1	Soil

5-48	JCM	3	1	Soil
5-6	JCM	3	1	Soil
5-8	JCM	3	1	Soil
6-3	JCM	3	2	Soil
6-4	JCM	3	2	Soil

Table A7: L-DOPA and 2,6-DCP oxidation for Rooibos isolate 2-8 over a 10-day period

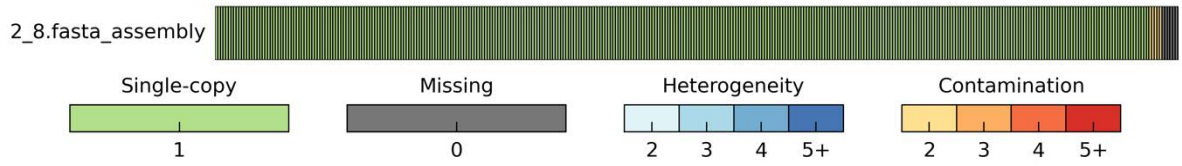
***Lentzea* sp. isolate 2-8**	**U/mL**			
	Day 3	Day 5	Day 7	Day 10
2,6-DCP oxidation	0.171294	0.03012	0	0.194022
L-DOPA oxidation	1.861333	0.034222	0	1.002778

Figure A3: CheckM output – isolate SF.14



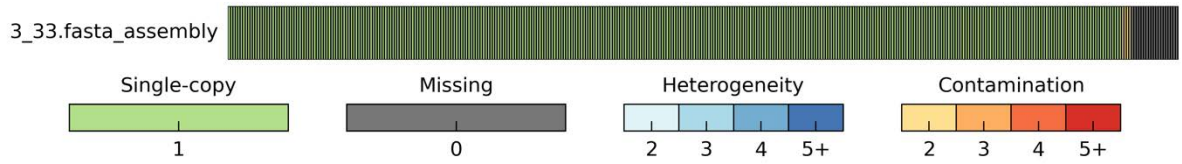
Bin Name	Marker Lineage	# Genomes	# Markers	# Marker Sets	0	1	2	3	4	5+	Completeness	Contamination
SF1_4_SPAdes.fasta_assembly	o__Actinomycetales	455	310	187	1	303	6	0	0	0	99.47	2.85

Figure A4: CheckM output – isolate 2-8



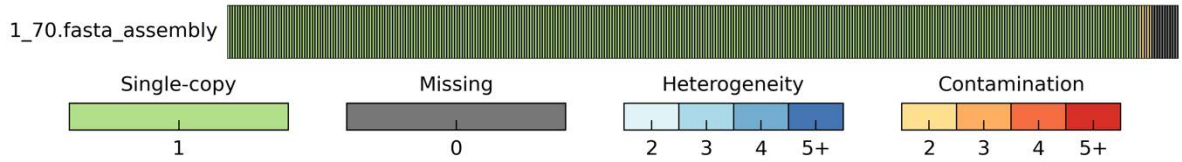
Bin Name	Marker Lineage	# Genomes	# Markers	# Marker Sets	0	1	2	3	4	5+	Completeness	Contamination
2_8.fasta_assembly	o__Actinomycetales	334	368	206	6	357	5	0	0	0	97.65	1.04

Figure A5: CheckM output – isolate 3-33



Bin Name	Marker Lineage	# Genomes	# Markers	# Marker Sets	0	1	2	3	4	5+	Completeness	Contamination
3_33.fasta_assembly	o__Actinomycetales	274	388	214	19	365	4	0	0	0	94.35	1.11

Figure A6: CheckM output – isolate 1-70



Bin Name	Marker Lineage	# Genomes	# Markers	# Marker Sets	0	1	2	3	4	5+	Completeness	Contamination
1_70.fasta_assembly	o__Actinomycetales	334	368	206	10	353	5	0	0	0	97.25	1.94

Figure A7: CheckM output – isolate 5-12

Bin Name	Marker Lineage	# Genomes	# Markers	# Marker Sets	0	1	2	3	4	5+	Completeness	Contamination
5_12.fasta_assembly	o__Actinomycetales	69	400	198	4	396	0	0	0	0	99.44	0.0

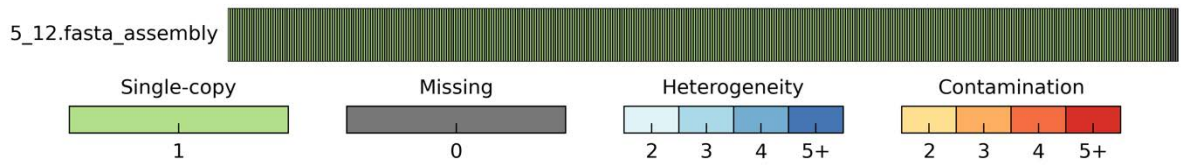


Table A8: Summary of Rooibos isolates using ABTS and 2,6-DMP exhibiting extracellular MCO activity above an arbitrary volumetric value of 0.05 U/mL

Isolate	U/mL	SD	Peak Day	Substrate	U/mL	SD	Peak Day	Substrate
3-33	0.32	0.01	3	ABTS				
1-60	0.25	0.09	10	ABTS				
6-4	0.24	0.00	10	ABTS				
1-8	0.21	0.07	5	ABTS	0.12	0.01	10	2,6-DMP
1-59	0.16	0.01	10	ABTS				
1-48	0.13	0.00	5	ABTS	0.10	0.01	10	2,6-DMP
1-50	0.13	0.03	5	ABTS	0.11	0.02	10	2,6-DMP
4-50	0.13	0.01	5	ABTS	0.10	0.01	10	2,6-DMP
4-77	0.13	0.02	5	ABTS	0.10	0.01	10	2,6-DMP
5-4	0.13	0.09	10	ABTS				
1-66	0.12	0.04	10	ABTS				
5-32	0.12	0.01	5	ABTS	0.11	0.00	10	2,6-DMP
5-33	0.12	0.01	5	ABTS	0.09	0.01	10	2,6-DMP
4-58	0.11	0.00	5	ABTS	0.11	0.01	10	2,6-DMP
3-34	0.11	0.03	10	ABTS				
3-8	0.10	0.01	3	ABTS				
1-61	0.09	0.00	5	ABTS				
4-30	0.07	0.01	10	ABTS				
4-54	0.07	0.01	10	ABTS				
5-12	0.06	0.00	3	ABTS	0.12	0.00	7	2,6-DMP
1-52					0.11	0.00	3	2,6-DMP
1-70					0.10	0.00	3	2,6-DMP
1-71					0.09	0.02	3	2,6-DMP
1-73					0.10	0.02	3	2,6-DMP
2-23					0.11	0.00	3	2,6-DMP
2-38					0.09	0.00	5	2,6-DMP
2-39					0.08	0.00	3	2,6-DMP
3-30					0.11	0.00	3	2,6-DMP
3-36					0.06	0.02	5	2,6-DMP
3-50					0.11	0.03	5	2,6-DMP
3-6					0.09	0.05	7	2,6-DMP
3-69					0.11	0.01	3	2,6-DMP
4-19					0.06	0.01	3	2,6-DMP
4-22					0.06	0.04	10	2,6-DMP
4-25					0.11	0.00	5	2,6-DMP
4-26					0.10	0.01	3	2,6-DMP
4-27					0.10	0.02	3	2,6-DMP
4-37					0.09	0.00	3	2,6-DMP
4-53					0.11	0.00	3	2,6-DMP
4-59					0.05	0.00	5	2,6-DMP
5-14					0.11	0.00	3	2,6-DMP
5-15					0.07	0.00	7	2,6-DMP
5-3					0.11	0.00	3	2,6-DMP
5-43					0.09	0.01	5	2,6-DMP
6-3					0.11	0.00	3	2,6-DMP

Table A9: Amino acid sequences for the chosen four MCOs identified in the genomes of four Rooibos isolates

Protein	Amino Acid Sequence
1-70_MCO	MSPVDRRRFLGLSGLAVVGGVASIGIGTGFSFGRVLDGSGEPGVLVRSGARLPQAFRSALPIPPV LRPTRSDTTTDYYEITQKAADVEYLPGLRTPSWTYNGSFPGPTLITRSGRRAVVTHRNELPRPVV VHLHGGHMPADSDGYPGDTILPRDGSSVSHDMSVPGNPVVGSRDYVYPGKQRAATLWYHDHSMGF TGATVYRGLAGFHLVRDDEDDALPLPKGDRDIPLMITDRSFAADGQFAYPSLAPNLTVPGVTEDH LNGVLGDVVLVNGAPWPALPVDRRRYRFRILNASNCRRYGLSLDPPPPDGGPAFTQIGSDGGLLP RPLTHESIDVAPAERFDVVLDFARYAPGTRVRLVNTLATDRTGEVMCFDVSDRTPRDTTAIPDEL SSVEYLDPRQAVRTREFLFQSKNGDPGWSINGEPYTPGTVLAHSRLGDLELWRFTSDVHHPVHVH LNHFQVTRRSNGGPGPYDAGWKDTVDLHPAQAMEVAIRFTDYPGRFVFHCHNLEHEDMGMMADFT TL
2-8_MCO	MRIRRVVTTRAGGVSRGSFESFNLGDHVGDDVEAVEANRVRLAEGIGLAPDRLVWMEQVHGRTVA TVDGPRAEPLEATDAVVTKRGGLGLVVLTADCVPVLLGDQEAGVVGAVHAGRVGARVGVVVEALK AMMALGAELERVEVLLGPAVCGECYEVPADMQRDVEKHLPGSASKSRKGTPALDLRAGLWNQLAS AGVGKIGVDPRCTFEEKDLFSHRRQAPTGRLASVVWVEP
SF1.4_MCO1	MSEHGDDRPVRARDLARRHVLSAGGALGLVAFTGLTTAHALARRPPRTGAALRSEVPLPPPFQVP LPLPSVLKPVGTAGGIDRYEITQRETTAEILPGVRTPLWTYGGTFPGPTIESRRGRPVTVRHRNE LPVPTVVHLHGGRTPAASDGYPTDLVLPKAWPGSAHGMGGMHMGGMSGMRDPRAAETRLVRDYTF PLDQRPTLLWYHDHRMDFTAPAIWRGLAGLHIVRDDAEDALGLPSGHRELPLMITDRAFGAGGHL RYPALDPSLRERPGVQEPYLAGVLGDVILVNGAPWPVHEVDAARYRLRVLNASNARHYDLEAVTD DGRRLDLVQVGADQGLLAAPVVHRSLPVAPAERYDLVVDFARVPVGGRVRIVNRLGSGRARDVMA FRVARKVRDGSRVPRVLSSDVPVWRRSEAVRVRDFSFRAGRMDGGHGWLIGGRPFDPARTDVTVR LGDVEVWRLVADVHHPVHLHLVGFRVLSRDGGPPPPHDAGLKDTVSLRPGEAVEIITRFDGYRGR YLFHCHNAEHEDMGMMANLEVV
SF1.4_MCO2	MELLDRRAMLRLSAGGAVAPVLAGRRRAAGAVPAAFSVPMTVPAELRPVRSTRDADHYRLVAAPG WAEILPGVRTPVLAYNGAFPGPTIRARAGRRAVVEYVNRLGEPTTVHLHGGDVWPEDDGGPMDLV APGTARRYRYPDRQRAATLWYHDHAHHLESEHVYRGLAGTYVLHDRHEAALGLPGGRYDVPIVLK DARFGADGGLVYIPGDFGGRTAILVNGRPQPYFRVSARKYRLRLLNAANMRFFNARLSTGDAFQQ IATDRGLLPAPLTTTELPLSPGERADIVVDFSRYAPGTRIVLRDGFSTHPATDAILRFDVCGRVR DGSRVPERLTALPALPPAAVTREFRLSMDERTGMGFINGRTFDAARIDFTVARGAAETWRIVNDN QVIEHNFHVHLADFRVLDRNGGPPLPTERGLKDTVMLPPGGSVRIRVSFDPRYCGDYVYHCHVID HSSMGMMGRFRVAGHR

Figure A8: InterPro summary for SF1.4_MCO1

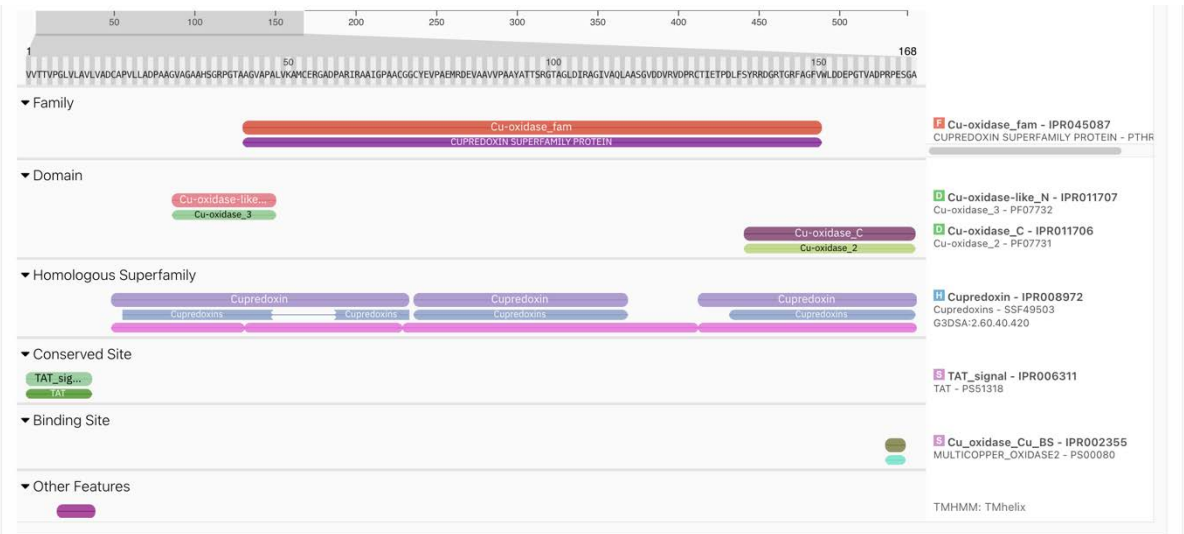


Figure A9: InterPro summary for SF1.4_MCO2

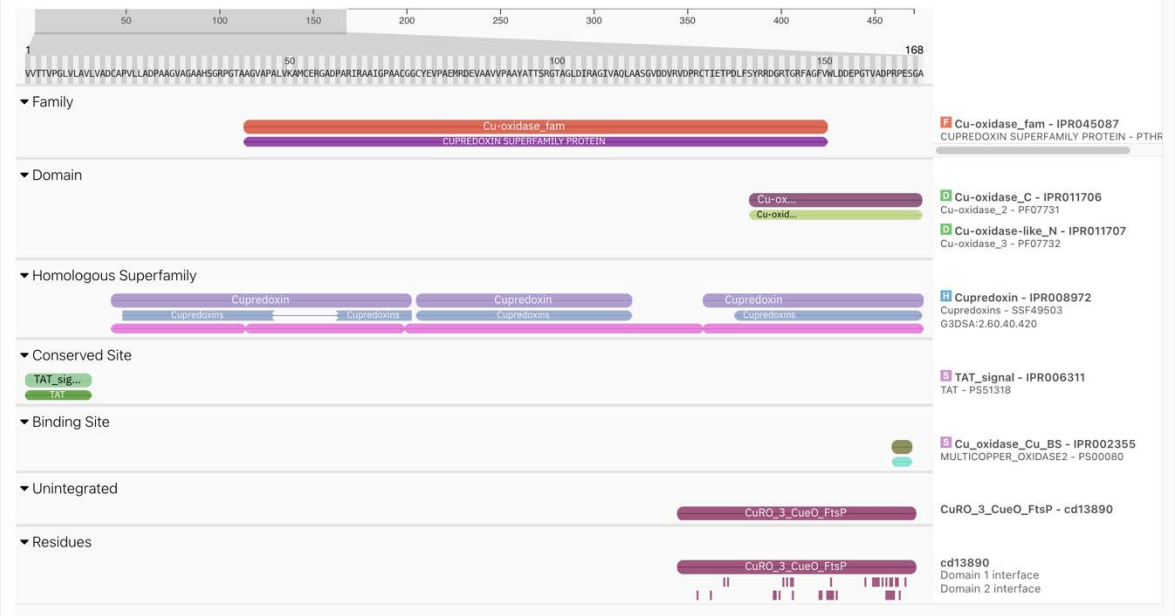


Table A10: pSORTb output for MCOs identified from the genomes of Rooibos isolates

```
SeqID: Actinokineospora_Isolate_1-70_MCO_001
  Analysis Report:
    CMSVM+            Unknown                    [No details]
    CWSVM+            Unknown                    [No details]
    CytoSVM+          Cytoplasmic                [No details]
    ECSVM+            Unknown                    [No details]
    ModHMM+           Unknown                    [1 internal helix found]
    Motif+            Unknown                    [No motifs found]
    Profile+          Unknown                    [No matches to profiles found]
    SCL-BLAST+        Cytoplasmic                [matched 16077697: spore coat protein (outer) [Bacillus
subtilis subsp. subtilis str. 168]]
    SCL-BLASTe+       Unknown                    [No matches against database]
    Signal+           Unknown                    [No signal peptide detected]
  Localization Scores:
    Cytoplasmic           9.97
    CytoplasmicMembrane   0.00
    Cellwall              0.01
    Extracellular         0.02
  Final Prediction:
    Cytoplasmic           9.97
  Secondary localization(s):
    Spore
--------------------------------------------------------------------------------
SeqID: Lentzea_Isolate_2-8_MPO
  Analysis Report:
    CMSVM+            Unknown                    [No details]
    CWSVM+            Unknown                    [No details]
    CytoSVM+          Cytoplasmic                [No details]
    ECSVM+            Unknown                    [No details]
    ModHMM+           Unknown                    [No internal helices found]
    Motif+            Unknown                    [No motifs found]
    Profile+          Unknown                    [No matches to profiles found]
    SCL-BLAST+        Unknown                    [No matches against database]
    SCL-BLASTe+       Unknown                    [No matches against database]
    Signal+           Unknown                    [No signal peptide detected]
  Localization Scores:
    Cytoplasmic           7.50
    CytoplasmicMembrane   1.15
    Cellwall              0.62
    Extracellular         0.73
  Final Prediction:
    Cytoplasmic           7.50
--------------------------------------------------------------------------------
SeqID: Actinomadura_SF1.4_MCO_001
  Analysis Report:
    CMSVM+            Unknown                    [No details]
    CWSVM+            Unknown                    [No details]
    CytoSVM+          Cytoplasmic                [No details]
    ECSVM+            Unknown                    [No details]
    ModHMM+           Unknown                    [1 internal helix found]
    Motif+            Unknown                    [No motifs found]
    Profile+          Unknown                    [No matches to profiles found]
    SCL-BLAST+        Cytoplasmic                [matched 16077697: spore coat protein (outer) [Bacillus
subtilis subsp. subtilis str. 168]]
```

```
    SCL-BLASTe+       Unknown                     [No matches against database]
    Signal+           Unknown                     [No signal peptide detected]
  Localization Scores:
    Cytoplasmic            9.97
    CytoplasmicMembrane    0.00
    Cellwall               0.01
    Extracellular          0.02
  Final Prediction:
    Cytoplasmic            9.97
  Secondary localization(s):
    Spore
--------------------------------------------------------------------------------
SeqID: Actinomadura_SF1.4_MCO_002
  Analysis Report:
    CMSVM+            Unknown                     [No details]
    CWSVM+            Unknown                     [No details]
    CytoSVM+          Cytoplasmic                 [No details]
    ECSVM+            Unknown                     [No details]
    ModHMM+           Unknown                     [No internal helices found]
    Motif+            Unknown                     [No motifs found]
    Profile+          Unknown                     [No matches to profiles found]
    SCL-BLAST+        Cytoplasmic                 [matched 16077697: spore coat protein (outer) [Bacillus
subtilis subsp. subtilis str. 168]]
    SCL-BLASTe+       Unknown                     [No matches against database]
    Signal+           Unknown                     [No signal peptide detected]
  Localization Scores:
    Cytoplasmic            9.97
    CytoplasmicMembrane    0.00
    Cellwall               0.01
    Extracellular          0.02
  Final Prediction:
    Cytoplasmic            9.97
  Secondary localization(s):
    Spore
--------------------------------------------------------------------------------
```

Table A11: Cloning primers used in this study to clone new MCOs

Cloning Primers	Sequence (5’ -> 3’)
SF1.4_MCO1_F	AGACACCATATGatggagctgctcgac
SF1.4_MCO1_R	AGATGGCTCGAGgcggtggccggcgac
SF1.4_MCO2_F	GACCTAGCGATCGCatgagtgagcacggc
SF1.4_MCO2_R	TGCCCAGTTTAAACtcagacgacctcgag
2-8-MPO_F	ATTCATATGgtgcgcattcgtcg
2-8_MPO_R	ATTAAGCTTtcacggctcaaccc

Table A12: GitHub links to Python packages used for EnzymeML parameter estimation.

Module	
sdRDM	https://github.com/JR-1991/software-driven-rdm
MTPHandler	https://github.com/FAIRChemistry/MTPHandler
CaliPytion	https://github.com/FAIRChemistry/CaliPytion
EnzymePynetics	https://github.com/haeussma/EnzymePynetics

APPENDIX B

List of isolation media used for the isolation of strains from the Rooibos environment:

All media prepared in 1L batches. pH was adjusted to pH 6.0.
Agar: in addition to the amounts below, 20 g of agar was added after pH adjustment.
All media were autoclaved at 121˚C, 15 psi, for 20-25 minutes.

	g/L
ISP medium no. 2	
Yeast extract	4.0
Glucose	4.0
Malt Extract	10.0
R2A	
Yeast extract	4.0
Peptone	1.0
Casamino acids	1.0
Glucose	1.0
Starch	1.0
Sodium tartrate	1.0
K_2HPO_4	0.6
$MgSO_4.7H_2O$	0.1
JCM medium no. 61	
Soluble starch	15.0
Yeast extract	4.0
K_2HPO_4	0.5
$MgSO_4.7H_2O$	0.5
Czapek Solution	
Sucrose	30.0
$NaNO_3$	2.0
K_2HPO_4	1.0
KCl	0.5
$MgSO_4.7H_2O$	0.5
$FeSO_4.7H_2O$	0.01
Starch-Casein-Nitrate	
Starch	10
Casein	0.3
KNO_3	2.0
$CaCO_3$	0.3
$MgSO_4.7H_2O$	0.05
$FeSO_4.7H_2O$	0.01

HV

Humic acid	1.0
$CaCO_3$	0.02
NaH_2PO_4	0.5
KCl	1.7
$MgSO_4.7H_2O$	0.5
$FeSO_4.7H_2O$	0.01

0.1% rooibos Extract

1.0 g of rooibos dust steeped in warm 1L dH_2O for 20 minutes and filtered with a size 102 coffee filter (House of Coffees). The extract was then made up to 1 L with dH_2O prior to adjustment of the pH and the addition of agar.

List of isolation media used for the transformations and expressions of constructs in *E. coli*

All media prepared in 1 L batches. pH was adjusted to pH 7.0.

Agar: in addition to the amounts below, 15 g of agar was added after pH adjustment.

All media were autoclaved at 121˚C, 15 psi, for 20-25 minutes.

SOC medium

2.0 g	tryptone
0.5 g	yeast extract
1 mL	1M NaCl
0.25 mL	1M KCl

After autoclaving, cool media and add:

1 mL	2M Mg^{2+} stock solution
1 mL	2M Glucose, filter-sterilised

Autoinduction medium (AI)

Component 1: ZY

10 g tryptone, 5 g yeast extract in 1L distilled water; autoclave. This can be kept at 4°C.

Component 2: 20× NPS

Add 6.6 g $(NH_4)_2SO_4$, 13.6 g KH_2PO_4, and 14.2 g Na_2HPO_4 to 90 mL distilled water. Dissolve and adjust the pH to 6.75. Once the pH is correct, make it up to 100 mL and autoclave.

Solution 1: 150× 5052

Add 25 g glycerol, 2.5 g glucose and 10 g lactose to 100 mL distilled water. Mix to dissolve and autoclave.

Solution 2: 1 M $MgSO_4$

Add 24.65 g of $MgSO_4.7H_2O$ to 100 mL distilled water. Filter-sterilize.

Component 1:	232 mL

Solution 1:	250 μL
Component 2:	5 mL
Solution 2:	12.5 mL
Antibiotic(s)	250 μL

2xYT Media

16.0 g	Tryptone
5.0 g	NaCl
5.0 g	Yeast extract

Luria Broth

10.0 g	Tryptone
5.0 g	Yeast Extract
10.0 g	NaCl
