INVESTIGATING THE OIL BIOREMEDIATION CAPACITIES OF THERMOPHILIC BACTERIA

by

Michael Mol

(788559)

Dissertation

Submitted in fulfilment of the requirements for the degree

Master of Science

in

Molecular and Cell Biology

in the Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa

Supervisor: Dr. Pieter De Maayer

Co-supervisor: Prof. Don A. Cowan

February 2020

DECLARATION

I declare that this thesis is my own, unaided work. It is being submitted for the Degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

_____

(Signature of candidate)

4th day of February 2020 at the University of the Witwatersrand

Abstract

Crude oil and petroleum derivatives are released, both inadvertently and actively, into vulnerable environments in vast amounts. As such the development and application of effective remediation approaches are major research imperatives. In particular, the development of environmentally friendly remediation approaches, such as bioremediation are focal points. One promising approach is the bioremediation of oil hydrocarbons using thermophilic bacteria. In Chapter 1 we have reviewed pertinent literature characterising crude oil and petroleum derivatives, their major environmental releases, remediation approaches and contextualised the application of thermophiles in bioremediation. In Chapter 2 we conducted plate- and liquid-based screening assays on a substantial collection of thermophilic bacteria, identifying numerous strains capable of degrading used-motor oil. Furthermore, we identified a number of thermophilic taxa which are capable of oil displacement, most likely linked to the production of biosurfactants. In Chapter 3 we further characterised the oil hydrocarbon degradation capacity of three strains belonging to the genus *Geobacillus*. This involved whole genome sequencing of these strains, as well as qualitative and quantitative characterisation of the hydrocarbon degradation capacities of the strains using 2D GC/MS analysis. One strain in particular, *Geobacillus stearothermophilus* Tok5A2 showed great promise in terms of its oil degradation activity. This strain possesses several aliphatic hydrocarbon degradation proteins, including three alkane monooxygenases, LadAα_{B23}, LadAβ_{B23} and LadB$_{B23}$ which may underlie this capacity.

Acknowledgements

I would like to acknowledge Dr. Pieter De Maayer for his continued guidance throughout the project, Professor Don A. Cowan for his guidance and provision of the bacterial strains used in the project, Dr. Maya Makatini and the University of the Witwatersrand Department of Chemistry for their assistance. Furthermore, I would like to thank my lab partner Thalia Yenketsammy and my wife Bronwyn Mol for all their assistance. I would also like to acknowledge the University of the Witwatersrand PMA, GDARD and CSIR-DST for funding.

List of figures

List of figures (Continued)

List of tables

List of Supplementary Tables:

Nomenclature

- 2D GC/MS Two-dimentioanl gas chromatography-mass spectrometry
- BTEX Benzene, toluene, ethyl benzene and xylene
- dDDH Digital DNA-DNA Hybridization
- GGDC Genome to genome distance calculator
- LB Luria Bertani
- MAFFT Multiple Alignment Fast Fourier Transform
- MAHs Monoaromatic hydrocarbons
- MFS Major Facilitator Superfamily
- mLB Modified Luria Bertani
- NA Nutrient agar
- PAHs Polyaromatic hydrocarbon
- pMMO Membrane bound methane monooxygenase
- RAST Rapid annotations using subststems technology
- sMMO Cytoplasmic methane monooxygenase
- TYGZ Type (strain) genome server

CONTENTS **PAGE**

CHAPTER ONE: CHARACTERISATION, USE AND REMEDIATION OF CRUDE OIL AND PETROLEUM DERIVATIVES

CHAPTER TWO: SCREENING FOR THERMOPHILIC PETROLEUM HYDROCARBON DEGRADATION

CHAPTER ONE

CHARACTERISATION, USE AND REMEDIATION OF CRUDE OIL AND PETROLEUM DERIVATIVES

1.1 Introduction

Crude oil and its derivatives serve as the largest contributor to global energy production and are dependent on a global supply and distribution network (Rubin, 2012). During the harvesting, distribution and use of crude oil, there is a continuous risk of inadvertent release of crude oil and its derivatives, often into environments that are hundreds of kilometres from the origin or destination of the substances (Akob *et al*., 2016). These spillages pose significant threats to both the health of affected environments and their inhabitants (Holliger *et al*., 1997). Consequently, remediation of contaminated environments is of crucial importance to ensure continuing environmental health and conservation of affected ecosystems (Hakima & Ian, 2017).

While there are numerous existing strategies for remediation of affected environments, they are often limited in terms of their efficacy and cost effectivity. Bioremediation strategies, which utilize bacteria, fungi, yeasts and plant species capable of degrading pollutants, have proven to be an effective means of remediation of environments affected by mass contamination events (Atlas, 2011). Bioremediation techniques are typically more environmentally friendly and cost effective than chemical- or physicochemical-based remediation techniques (Das & Chandram, 2011). Thus, the drive for the application of bioremediation techniques is ever increasing (Lim *et al*., 2016). Despite this, the depth of knowledge regarding species capable of petroleum oil degradation is limited, particularly regarding the capabilities of thermophilic bacteria, which generally present more traits conducive to rapid, effective and efficient removal of pollutants (Mehta *et al*., 2016). Further elucidation of the capabilities and application of these species may present more efficient and cost-effective means of oil spill bioremediation (Lim *et al*., 2016).

The aim of this thesis was to further elucidate the petroleum hydrocarbon degradation capacities of thermophilic bacteria. This was approached through extensive review of relevant literature (Chapter 1), screening a collection of aerobic thermophilic bacteria for the capacity to degrade hydrocarbons contained in used motor oil (Chapter 2) and molecular and genomic characterisation of hydrocarbon degradation observed by selected strains (Chapter 3).

Here we have reviewed relevant literature relating to the topics of crude oil composition, use, environmental contamination and contaminant mitigation techniques.

1.2 Crude oil

Crude oil or petroleum and its refined derivatives, gasoline and diesel are internationally dominant and essential energy sources, accounting for almost a third of global energy use (Rubin, 2012). In 2017, crude oil and oil-based products accounted for 32% of all fuels used for global energy production, registering as the largest contributor to global energy generation, ahead of the next largest contributor coal, accounting for 27% (Enerdata, 2018). Given the extensive use of crude oil-based products globally, the refinement of its raw state into its constituting fractions for further processing is key. Crude oils have highly variable composition, but are predominantly composed of four categories of molecules, namely aliphatic and aromatic hydrocarbons, resins and asphaltenes (Table 1.1) (Varjani, 2017).

1.2.1 Aliphatic hydrocarbons

Predominant aliphatics found in crude oil include saturated hydrocarbons, namely paraffins and cycloalkanes/naphthenes (Riazi, 2005; Varjani, 2017). Paraffins typically possess chain lengths ranging C_1–C_{40} and may represent up to 32,5% of the total volume of the oil (Rehan *et al*., 2016). Cycloalkanes also known as naphthenes may represent up to 60% of crude oil volumes and along with paraffins are stable compounds, capable of persisting in the environment for extended periods (Riazi, 2005).

1.2.2 Aromatic hydrocarbons

Aromatic hydrocarbons are compounds which contain aromatic ring structures, being either monocyclic aromatic hydrocarbons (MAHs) or polycyclic aromatic hydrocarbons (PAHs) (Varjani, 2017). MAHs present four main compounds of interest with respect to crude oil usage, these are benzene, toluene, ethylbenzene and xylene (BTEX) (Yadav *et al*., 2012). Along with being essential compounds involved in automotive and industrial settings, they also present environmental and

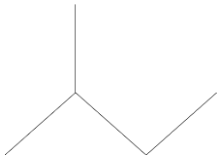
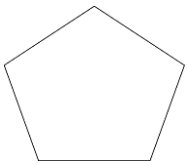
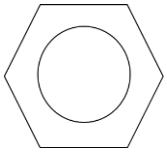
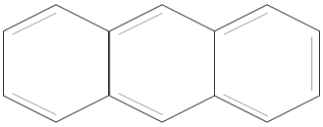
health risks. These compounds are of great concern as they are highly carcinogenic and toxic, even in low concentrations (Jiang *et al*., 2015). These compounds are relatively water soluble, which facilitates rapid and extensive dispersion in subsurface and marine environments (Yadav *et al*., 2012; Jaggi *et al*., 2017). PAHs contain at least two cyclic benzene rings (Abdel-Shafy & Mansour, 2016). Due to their cyclic structures, these compounds decrease in water solubility as their ring numbers increase, resulting in either residue or sediment formation in aqueous environments (Sun *et al*., 2016). This behaviour facilitates their persistence in these environments, which is problematic due to their carcinogenic and toxic attributes (Varjani, 2017).

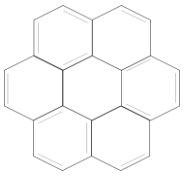
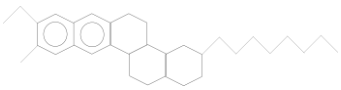
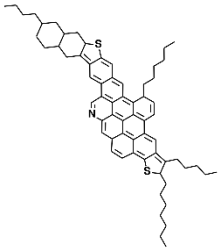
1.2.3 Resins and asphaltenes

Resins are compounds comprised of a polar end group attached to a long alkane tail. These end groups generally contain one or several aromatic or naphthenic rings, as well as sulphur, nitrogen or oxygen atoms and trace metals such as nickel, vanadium, copper and iron (Demirbas & Taylan, 2016). These polar groups are generally attached to long branched paraffinic tails (Demirbas & Taylan, 2016; Varjani, 2017). While little work has been conducted on the nature of crude oil resins, their structures however, are noted to be similar to asphaltene compounds, albeit of much lower molecular weight (Demirbas & Taylan, 2016).

Asphaltene compounds are larger and more complex compounds, composed of numerous condensed aromatic structures, with variable branching groups, sulphur, nitrogen and oxygen atoms and metals such as nickel, vanadium and iron (Demirbas & Taylan, 2016; Varjani, 2017). These compounds possess the greatest fraction of heteroatoms as well as some of the most toxic and persistent classes of compounds in crude oils (Demirbas & Taylan, 2016). They include esters, fatty acids, ketones, phenols and porphyrins, which produce the majority of the dark colouration typical of crude oils (Das & Chandran, 2011; Demirbas & Taylan, 2016). Together with resins, asphaltenes play crucial roles in stabilising crude oil, directly affecting particulate sedimentation and deposition in the environment (Demirbas & Taylan, 2016).

Table 1.1: Summary of hydrocarbon contents identified in crude oils

Hydrocarbon	Group	Structure	Qualities
Paraffins	Paraffins - Straight chained alkanes	 Pentane	C_1-C_{40} – Four classes of alkanes 1: Gaseous alkanes C_1-C_7 2: Low molecular weight hydrocarbons C_{8-16}. 3: Medium molecular weight hydrocarbons C_{17}-C_{28} 4: High molecular weight hydrocarbons >C_{28}
	Isoparaffins- Branched chain alkanes	 Isopentane	
	Cycloalkanes- Cyclic saturated hydrocarbons	 Cyclopentane	
Aromatics	MAHs benzene, toluene, ethylbenzene and xylenes	 Benzene	Toxic, carcinogenic, water soluble
	PAHs Low molecular weight	 Anthracene	2-6 benzene rings

Hydrocarbon	Group	Structure	Qualities
	PAHs High Molecular Weight	 Coronene	>6 benzene rings
Resins	Long chain alkane with polar end group	 Theoretical Resin	Highly polar end groups – ≥ 1 aromatic/ naphthenic rings
Asphaltenes	Numerous condensed aromatic structures variable branching groups	 Theoretical Asphaltene	Include - esters, fatty acids, ketones, phenols and porphyrins Variable branches - S, N, O, Ni, V, Fe

1.3 Refined crude oil fractions

Raw crude oils are rarely used without any form of processing or refinement (Speight, 2015). The different fractions of crude oils produced through refinement are channelled towards numerous and varied purposes (Riazi, 2005; Speight, 2015). However, general classification of these products is allocated as either fuel or non-fuel petroleum products (Riazi, 2005). Fuel based petroleum products include still gases ($>C_4$), liquified petroleum gases ($>C_4$), gasoline ($C_{5\text{-}12}$), naphtha ($C_{7\text{-}17}$) and kerosene ($C_{8\text{-}18}$) and generally comprise the lighter fractions of crude oil, which see greater domestic and commercial use (Speight, 2014; Speight, 2015). Heavier fraction fuel oils include diesel ($C_{14\text{-}20}$) and heavy fuel oils ($C_{20\text{-}70}$), comprise more complex compounds and see greater industrial use (Riazi, 2005; Speight, 2014). Non-fuel petroleum products are typically composed of mixed length paraffins and cycloalkanes ($C_{20\text{-}50}$) and are employed primarily in automotive and industrial applications as lubricating oils (Speight, 2014; Speight, 2015).

1.4. Petroleum spills and releases

As a result of the numerous roles crude oil derivatives assume in day to day life, their production, distribution and use across the globe are extensive. A by-product of this extensive use and distribution is the intentional and unintentional release of these compounds and their derivatives into the environment (Fathepure, 2014). While release of these substances into the environment is often focused on during large scale spills, numerous small-scale incidences account for the majority of the volume of petroleum products released annually. It has been estimated that large scale release events account for only 10% of annual release volumes, with the remaining amount accounting for cumulative small release events (Fathepure, 2014; Ivshina *et al*., 2015).

Upon spillage, the recalcitrant compounds in petroleum derivatives accumulate and persist in aquatic and soil environments, with marine environments serving as the eventual pool for these substances (Varjani, 2017). These substances pose numerous hazards to the systems they are deposited in and pass through, affecting both the health of the environment and the health of residing organisms (Tang *et al*., 2011). Many of the crude oil compounds discussed above have been identified as potential carcinogens, toxins and mutagens (Zhang *et al*., 1998; Abdel-Shafy & Mansour, 2016), having severe adverse effects on organisms inhabiting contaminated environments at both the macroscopic and microscopic level (Peterson, 2001; Jernelov, 2010).

Macroscopic effects include events such as smothering or entrapment of animals. Furthermore, petroleum compounds may contribute to extreme instances of environmental damage by lowering the environmental capacity for hosting organisms. This is accomplished through influencing the conditions and properties of the environment (Jernelov, 2010). These coupled with microscopic impacts including changes in soil permeability, water composition, light exposure and absorption, gaseous exchange, food web dynamics and the complex interactions between organisms and the environments they inhabit greatly alter natural systems (Jernelov, 2010; Gong *et al*., 2014; Vengosh, 2016).

The range of substances with varying properties and chemistries in crude oil and petroleum derivatives result in variable behaviour of spilled compounds in the environment (Gong *et al*., 2014). In aquatic environments where petroleum spills or accumulations are more prominent, lighter density compounds tend to locate as a layer at the surface, undergo emulsion and dissolution and dispersion through the marine environment (Ivshina *et al*., 2015). Heavier density compounds, typically the persistent complex long-chained resins and asphaltenes, tend to form emulsions in underwater plumes, aggregates and sediments, either as particulate mass in the environment or on the floor of the environment (Gong *et al*., 2014; Ivshina *et al*., 2015).

Petroleum compounds that form the various accumulations in marine environments impact the properties of the environment differentially. Lighter particles that form surface layer slicks tend to cause several major problems. These include, restricting light passage to lower zones in the environment, which may also coat surface dwelling organisms or be ingested causing harm to the organisms (USEPA, 2006; Jernelov, 2010; Gong *et al*., 2014). The passage of sunlight through marine water columns is intrinsically linked to the profiles of microbial communities found in the environment (Ruiz-Gonzales *et al*., 2013). The light absorbing properties of the compounds found in crude oils, and the alterations in the communities of microbiota in contaminated environments has the potential to greatly affect the biotic and abiotic profiles of the environment (Ruiz-Gonzales *et al*., 2013; Bacosa *et al*., 2015). Surface oil slicks are also most notably identified to have severe adverse effects on both seabirds and marine mammals such as whales, seals and otters (Peterson *et al*., 2003; Jernelov, 2010). Oil slicks are noted to cause entrapment, smothering and coating of seabirds, ultimately resulting in induced poor health and death (National Research Council, 2003; Peterson *et al*., 2003). Marine mammals such as whales, seals and otters which interact with surface and subsurface contaminated waters, are at risk of suffocation, poisoning and even induction of hypothermia from disruption of their natural insulation (Baker, 2001; National Research Council, 2003; Peterson, *et al.,* 2003).

Crude oil compounds are also prone to undergo partial dissolution, exposing organisms throughout the water column to chronic exposure to the compounds resulting in ingestion, coating or accumulation in the gills of marine animals (Baker, 2001; National Research Council, 2003; USEPA, 2006). Greater density compounds or complex compound mixtures comprising less soluble and degradable components tend to accumulate as aggregates on particulate matter in the water column or as impermeable sediments on the floor of affected environments (USEPA, 2006; Jernelov, 2010; Gong *et al*., 2014; Ivshina *et al*., 2015). Suspended aggregates are capable of being dispersed over large distances via natural marine currents and pose the risk of spreading contamination over extended distances around the initial spill location (National Research Council, 2003). Furthermore, the compounds forming the sediment may also influence microbial metabolic processes in the subsurface and surrounding areas of the sediment (Jonker *et al*., 2006). The toxicity of the sediment compounds may cause the eradication of numerous species of seabed residing organisms. Changes in these communities may result in long term changes in community composition, metabolic profiles and conditions of the affected environments. These alterations in the various profiles of the environmental microbial activity may cause further damage to water, seabed and surrounding environmental health, hosting capacity and properties (National Research Council, 2003; Cravo-Laureau & Duran, 2014; Hakima & Ian, 2017).

1.4.1 Small- and Large-scale contamination events

Small-scale contamination events

While large scale spills and releases of crude oils and petroleum products are typically better reported and documented, they comprise only a minor fraction of annual discharges into the environment (Ivshina *et al*., 2015). Small scale spill events are typically localised to sources of origin of crude oils, i.e. drill or pump points and as such do not pose as pressing threats as large-scale contamination events which release excessive quantities of contaminants into the environment at one time point. Due to the scale and media response to large scale release events, the latter have taken precedence in remediation and management efforts (Baker, 2001; Moroni *et al*., 2019).

Large-scale contamination events

Large-scale contamination events noted to have caused massive environmental impacts and attracted global attention include predominantly large shipping accidents or extraction point spills (Jernelov, 2010; Moroni *et al*., 2019). Key examples of such events include the sinking of the Exxon Valdez oil tanker and the BP Deepwater Horizon drill rig blowout.

The Exxon Valdez spill. The sinking of the Exxon Valdez in 1989 in Prince William Sound released an approximate 42 million litres of heavy crude oil into surface waters of the Gulf of Alaska. This spill resulted in oil covering a combined ~2,000 km of shoreline of the Prince William Sound and Gulf of Alaska coastlines (Atlas, 2011). The visible short-term impact on wildlife by this spill included the deaths of an estimated 300 seals, 1,000 – 2,800 sea otters and 250,000 seabirds. The long-term impacts of this spill ranged from a general increase in mortality across organisms in the affected area and decreased population recovery, health, growth and reproductivity of numerous species in the surrounding area (Peterson *et al*., 2003; Esler *et al*., 2018). The persistence of the toxic compounds released from the spill remain to date in beach and seabed sediments, sediment associated environments and the species inhabiting them. The persistence of the sedimented compounds have acted to facilitate chronic exposure of numerous species to the compounds, prolonging the effects of the spill (Peterson, 2001; Atlas, 2011; Esler *et al*., 2018).

BP Deepwater Horizon drill rig blowout. The blowout of the BP Deepwater Horizon drilling platform in 2010 into the Gulf of Mexico, released ~780 million litres of crude oil into the environment. This oil was released 1,500m below sea-level into the open ocean where it passed up to the surface via a thick oil plume. This oil collected as a surface slick covering over 110,000 km^2 of surface water, a large deep-water plume extending up to 400 km from the spill site, heavy sedimentation on the seabed and along approximately 2,100 km of the nearby shoreline (Atlas, 2011). The effects of this massive spill on the environment were extensive, impacting organisms throughout the food web, including plankton, invertebrates, marine mammals, fish, sea turtles and seabirds both in the short- and

long-term periods after the spill. Much like the Exxon Valdez spill the effects of residual contaminants persisting in the affected environments continue to negatively affect a broad spectrum of the organisms in the affected area (Beyer *et al*., 2016). Following the clean-up, remediation and management of the spill, an estimated 26% (~202 million litres) of the total volume of oil spilled remained in the environment (Atlas, 2011). Thus, it is apparent that the need for rapid and effective means of managing and remediating the effects of such spills in the environment is required to effectively preserve the diversity and health of the environment.

1.5 Petroleum hydrocarbon remediation practices

A broad range of techniques have been developed and employed in order to try and mitigate or diminish the effects of oil spills on the environment. The choice of approach is largely dependent on the conditions of the contamination site, whether contamination is soil- or water-bound and the environmental conditions of the contamination site (Ivshina *et al*., 2015). Approaches include thermal, physicochemical, and bioremediation strategies (Ivshina *et al*., 2015).

1.5.1 Thermal remediation strategies

Thermal remediation strategies involve the extreme heating of petroleum contaminated substrates, to either directly or indirectly assist their removal from the contaminated environment (Lim *et al*., 2016). Thermal strategies include soil-based approaches such as incineration, thermal desorption and microwave heating and marine/aquatic approaches such as *in situ* burning (Abbasian *et al*., 2015; Ivshina *et al*., 2015).

Incineration approaches involve the controlled heating of oil contaminated soils to temperatures above 800°C, whereby complete or near complete removal of contaminating substances is achieved (Abbasian *et al*., 2015). This technique is predominantly applied to large soil bound oil spills due to its speed and efficiency compared to other techniques. The main drawbacks of this approach are the large amount of pollution, which is released into the air, the very poor cost effectiveness

of the approach and the nutrient poor condition the environment is left in post incineration (Vidonish *et al*., 2016).

In situ burning is used as a rapid response or emergency response to large scale oil spills, or oil spills that occur in heavily iced areas such as in the arctic circle (Ivshina *et al*., 2015). This response is a comparatively rapid, simple and effective means of removing oil from marine environments that are likely to not respond well to alternative remediation strategies or require immediate bulk removal of the contaminating oil (Buist, 2003; Ivshina *et al*., 2015). The effectiveness of this approach is dependent on a range of factors including the type, age, quantity and behaviour of the spilt oil in the environment as well as the environmental conditions present at the spill site (Buist, 2003). While this technique is rapid and effective given the correct conditions, its primary drawback is the resultant air pollution created by the burning oil. However, for oil spills not in immediate proximity to vulnerable communities the effect of allowing the oil to remain in the environment more adversely affects the environment (Ivshina *et al*., 2015). Along with this problem, lesser concerns include the ignition of slicks and the associated hazards of these fires (Buist, 2003; Ivshina *et al*., 2015).

Thermal desorption and microwave heating rely on the processes of inducing degradation and volatilization of contaminating compounds in contaminated soils by either microwave or alternative sources of heating. Compounds are either degraded or volatilized in the process and are then extracted using sweep gases and further processed and degraded (Abbasian *et al*., 2015). These techniques facilitate fast and effective removal of contaminating oil compounds with minimal air pollution, however present potential health and safety concerns during operation due to compound volatilisation and the use of microwave radiation (Abbasian *et al*., 2015; Vidonish *et al*., 2016).

1.5.2 Physicochemical and chemical remediation strategies

Approaches towards physicochemical and chemical remediation strategies are governed by the environment the strategy is applied to. Terrestrial or marine/aquatic based contaminant remediation require vastly different techniques which are aimed

to maximize either contaminant recovery or contaminant removal from the afflicted environment.

Terrestrial remediation

Recovery of soil bound oil contaminants may be achieved through numerous methods, including solvent extraction/soil washing, soil vapor extraction, flotation and ultrasonication (Ivshina *et al*., 2015). However, the most prominent and widely applied recovery techniques are solvent extraction/soil washing and soil vapour extraction (Kuppusamy *et al*., 2016).

Solvent based extraction/soil washing involves the application of organic or non-organic solvents and flushing agents to contaminated soils. This facilitates partitioning of contaminants from the substrate, allowing for removal of solid and liquid contaminants (Kuppusamy *et al*., 2016). This technique has been developed to include surfactants or surfactant producing bacteria, which facilitate greater contaminant recovery (Kavitha *et al*., 2014). This technique is a highly effective recovery method, with low small-scale cost and generally low eco-toxicity. However, its large-scale application cost and the potential for secondary pollutant introduction are major hindrances to its application (Lim *et al*., 2016).

While solvent based extraction/soil washing targets predominantly solid and liquid phase contaminants, soil vapour extraction relies on the extraction/venting/ vacuuming of volatile contaminants out of the soil environment (Lim *et al*., 2016). Application of this technique is restricted due its environmental applicability. However, under ideal conditions, it is highly effective and cost efficient. While highly limited in isolated use, integrated approaches incorporating this technique have proven to be highly effective (Kuppusamy *et al*., 2016).

When recovery of oil contaminants from terrestrial environments is not feasible and removal is required, environmentally harsher techniques are applied. The most prominent remediation strategy employed in terrestrial contamination events are chemical oxidation approaches. Commonly applied treatments include Fenton's reagent, ozone and persulfate (Lim *et al*., 2016).

Fentons reagent has proved a viable mechanism for oil contaminant degradation (Ferguson *et al*., 2004; Lim *et al*., 2016; Usman *et al*., 2016). It involves the interaction of hydrogen peroxide with environmental soluble iron compounds, generating (OH^-) free radicals. These oxidize oil compounds facilitating their conversion into less harmful organic compounds (Lim *et al*., 2016). While highly efficient and effective, the effects of deteriorating soil quality and microbial hosting capacity, leave room for improvement of the technique (Lim *et al*., 2016). Further advancement is required to optimise application of chemical oxidation practices (Usman *et al*., 2016).

Marine/Aquatic remediation

Physical recovery of oil contaminants from marine environments generally involve concentration of the oil followed by separation of emulsified oil from contaminated water (Ivshina *et al*., 2015). The simplest and most commonly deployed mechanism for concentrating and preventing spreading of spilled oil is achieved using booms. Use of booms serve as an environmentally friendly and non-invasive method compared to others. The primary drawbacks of boom application are the slow rate of application and its susceptibility to environmental conditions such as turbulent or fast flowing water conditions (Ivshina *et al*., 2015; Hoang *et al*., 2018). Following concentration of spills in the affected environment, separation techniques generally involve oil-water skimming or use of absorbent or adsorbent materials or a combination of both. Various skimming operations such as oleophilic, weir and vacuum based techniques have been developed, each with advantages and disadvantages. General advantages include their minor invasiveness and lower environmental impact than chemical treatments or *in situ* burning. Disadvantages include their limited throughput before blockage, by debris or aggregation of oil coated particles, their labour intensiveness and particularly for sorbents, their limited use for small scale applications (Ivshina *et al*., 2015; Nyankson *et al*., 2016; Hoang *et al*., 2018).

There are a range of chemical treatments that are applied to water-based oil spills which affect the various physicochemical properties of the contaminating oil with

its environment. However, the two primary classes of chemicals applied are de-emulsifiers and dispersants (Ivshina *et al*., 2015).

De-emulsifiers are combinations of solvents and surfactants that reduce the ability of oils to form emulsions. The primary driver for use of these substances is to facilitate more efficient oil recovery from the spill site through greater separation of contaminating oil and environmental water. (Tamis *et al*., 2011; Ivshina *et al*., 2015). These substances are of particular importance for spills where the environment is turbulent, or the oil is released at the sea-floor and travels through the water column, facilitating emulsion formation (Tamis *et al*., 2011). The drawbacks of these substances lie in their toxicity to some marine organisms, their potential to facilitate oil passage through the water column in more concentrated volumes and their variability in effectiveness depending on oil type and environmental conditions (Tamis *et al*., 2011; Ivshina *et al*., 2015).

Dispersants are typically substances containing a solvent and a surfactant that are used to improve the miscibility of oils in water. Applications of dispersants are aimed at reducing interfacial tensions between oil and water and inducing the formation of small oil droplets in the water (Tamis *et al*., 2011; Ivshina *et al*., 2015). This facilitates more pronounced dispersion, susceptibility to degradation and lesser surface slick impacts of oils (Gong *et al*., 2014). These effects further assist in minimising harm to surface dwelling seabirds and mammals, as well as minimising surface bound transport of large volumes of oil into coastal areas where they are likely to further damage these ecosystems (Tamis *et al*., 2011). There are several drawbacks of dispersant usage covering several areas of concern. Application of dispersants is only effective on light and non-weathered oils and may only be applied effectively in areas where high dilution rates of the oil droplets are observed, largely limiting their application (Tamis *et al*., 2011). Some substances used in commercial dispersant applications have also been noted to be carcinogenic, toxic to humans and present unknown impacts on marine dwelling communities (Gong *et al*., 2014; Seidel *et al*., 2016). Along with this, the health effects of fine chemically dispersed oil droplets in marine habitats on their resident communities are also unclear (Seidel *et al*., 2016). As such their application is still not fully supported.

1.5.3 Bioremediation strategies

Due to the barriers encountered in employing the various aforementioned oil spill management and remediation techniques, interest in alternative, cost effective, efficient and environmentally friendly means of remediating oil spills has risen. The predominant focus of this movement has been the development and application of bioremediation approaches to oil spills (Chandra *et al*., 2013).

Bioremediation based approaches apply the principle of transformation or mineralisation of contaminant hydrocarbons through natural degradation pathways of microbes and plants (Vidali, 2001; Lim *et al*., 2016). The predominant by-products of these pathways include carbon dioxide, water or microbial cell constituents, organic matter, metabolic products or less complex organic and inorganic molecules (Margesin & Schinner, 2001). As such along with these being more cost effective, and less invasive means of petroleum hydrocarbon removal than other methods, they are the pinnacle of environmentally friendly remediation approaches (Dzionek *et al*., 2016).

The application of bioremediation techniques is determined by the nature of the contaminant and the contaminated environment itself (Gong *et al*., 2014). There are two categories of bioremediation practices, *ex situ* and *in situ,* conducted both terrestrially and in marine/aquatic environments. *Ex situ* bioremediation approaches employ the practice of removing contaminated substrate or media from the environment and transporting it to a location apart from the environment affected to be treated (USEPA, 2006). *In situ* bioremediation methodologies, however, involve local remediation of contaminated substrates and environments (Azubuike *et al*., 2016). The main advantages of *in situ* techniques are their application at the site of contamination, limiting further dispersion or distribution of contaminant and contaminated substrate as well as the relative ease of application and lesser cost as compared to *ex situ* techniques. The central disadvantage of these approaches however is the lack of control over the broader interactions of the conditions in the contaminated environment, limiting full regulation of the degradative process (Dzionek *et al*., 2016). Both approaches include techniques which induce rapid and

enhanced biodegradation of naturally occurring and introduced organisms (Vidali, 2001).

Terrestrial bioremediation strategies

Terrestrial bioremediation approaches incorporate a wide range of practices, both *in situ* and *ex situ*. There are numerous techniques which have been developed, with several having been used extensively in common practice. These include land farming, composting, biopiling, bioventing and biosparging, bioaugmentation and phytoremediation.

Land farming involves excavation, relocation and spreading of contaminated soil onto healthy soil, followed by application of traditional agricultural practices on the land used for remediation (Vidali, 2001; Kuppusamy *et al*., 2016). This involves tilling, fertilization, bulking and irrigation of the soil, promoting intrinsic microbial degradation of the contaminating compounds (Vidali, 2001). The advantages of this approach are its high effectivity and cost effectivity over extended time periods. However, extended operation periods result in extremely slow turnover rates of relatively low contaminant concentrations (Dzionek *et al*., 2016).

Composting is a more advanced approach compared to land farming, which involves the mixing of contaminated soils with amendments of sewage and agricultural and municipal waste (Dzionek *et al*., 2016). Composting presents nutrient and mineral rich substrate and fluctuation between mesophilic and thermophilic conditions with an overall increase in average temperature to 50°C (Hesnawi & McCartney, 2006; Kuppusamy *et al*., 2016). These conditions facilitate shifts in the microbial communities and microbial metabolisms in the compost environment, facilitating more efficient contaminant degradation (Hesnawi & McCartney, 2006). Composting presents overall positive appeal towards terrestrial bioremediation, its only disadvantage being its inefficiency in processing time compared to other approaches (Macauleay & Rees, 2014).

Biopiles are an altered approach to composting practices, which rely on the same principle of mineralisation of contaminant oil compounds by microbial communities, but are supplemented with aeration (Vidali, 2001; Jabbar & Kadhim,

2018). Biopiles are constructed as contained cells which prevent loss of contaminating compounds that are prone to volatilization or leaching (Vidali, 2001). Similarly, bioventing and biosparging rely on the natural activity of resident oil degrading organisms in the soil environment (Lim *et al*., 2016). In bioventing, a constant feed of oxygen is supplied at low pressure to contaminated soil such that growth of aerobic bacteria and subsequent contaminant degradation is accelerated. Biosparging applies the same practice but is conducted on saturated soils under the water table and at high pressures (Macauleay & Rees, 2014). While these systems are more efficient and less invasive than landfarming or composting, they are more cost intensive (Azubuike *et al*., 2016). These strategies have several pitfalls however, as they are faced with the disadvantages of numerous uncontrolled variables presented by the affected environment (Lim *et al*., 2016). Furthermore, poor environmental controllability may impact highly on the efficacy of these techniques (Lim *et al*., 2016).

While the previous techniques rely on intrinsic microbial communities for contaminant degradation, bioaugmentation involves the addition of foreign microbial species to the affected environment (Azubuike *et al*., 2016). These species are often genetically engineered or selected from the contaminated environment, with the intent of utilizing enhanced or specific metabolic pathways to increase the rate of contaminant biodegradation (Perelo, 2010). Bioaugmentation approaches supplement the capacity of intrinsic microbial communities for degrading oil pollutants (Adams *et al*., 2015). This may impact either the community's response rate to the contaminant or their ability to degrade the contaminant (Dzionek *et al*., 2016). This approach presents a relatively cost effective, minimally intrusive and environmentally friendly means of remediation. However, concerns about the efficacy of bioaugmentation include the capacity of the foreign organisms to survive in the polluted environment they are introduced to and their ability to compete with native organisms (Adams *et al*., 2015; Dzionek *et al*., 2016). Furthermore, inconsistencies displayed by field trials of bioaugmentation as a sole means of remediation represent a major prohibiting factor in its implementation (Macauleay & Rees, 2014; Azubuike *et al*., 2016).

Whilst above mentioned techniques have included use of exclusively microbial communities, phytoremediation makes use of plant and microbial communities. Either through direct interaction or through interplay between microbial, biological or inorganic processes, phytoremediation has been used as an effective means of remediation (Azubuike *et al*., 2016). As such, phytoremediation is employed in conjunction with other forms of bioremediation such as bioaugmentation or through construction of artificial wetlands, where interactions which benefit contaminant removal are supported (USEPA, 2006; Lim *et al*., 2016). While these approaches may be large in scale, they are generally cost effective, long term, and environmentally friendly (Azubuike *et al*., 2016). However, limitations restricting implementation include the extended turnover time of these operations, slow plant growth rates and resultant modes of action, limited zones of activity and the effects of concentration of pollutants on plant viability and activity (USEPA, 2006; Azubuike *et al*., 2016; Lim *et al*., 2016).

Marine/Aquatic and groundwater bioremediation strategies

While marine/aquatic and groundwater bioremediation strategies target a different phase of contaminated substrates to terrestrial remediation approaches, the techniques and principles applied are largely similar. As such most approaches are highly similar to terrestrial. There are, however, several common additional practices employed. (Dzionek *et al*., 2016). These include *ex situ* bioreactor processing and *in situ* biostimulation.

Bioreactor processing uses a reaction process that occurs in a reactor chamber with highly controlled parameters, with selected microbes (Vidali, 2001). Bioreactor remediation of contaminated water, sludge or sediment involves the extraction or pumping of the contaminated substrate into an appropriate bioreactor unit. Factors influencing the type of reactor used include the nature of the microbes (aerobic or anaerobic) used, the reaction conditions required and the nature of the contaminated substrate (Azubuike *et al*., 2016). Advantages of bioreactor approaches include their rapidity, controllability and effectivity (Vidali, 2001). There are however, several hinderances of employing bioreactors, including cost, complications associated with substrate transportation and bioreactor setup optimization (Vidali,

2001; Azubuike *et al*., 2016). Furthermore, the use of bioreactors is fairly limited to semi-solid- and solid-state contaminated substrates, limiting its application in surface and waterbody bound spills (Macauleay & Rees, 2014).

In situ biostimulation involves the delivery of essential nutrients to environments contaminated by petroleum compounds to stimulate intrinsic microbial communities to elevate their degradation of the compounds (Macauleay & Rees, 2014). The main groups of nutrients delivered are comprised of nitrogen and phosphorous rich compounds which are essential for bacterial growth. This approach shares similar rationales with terrestrial approaches such as composting whereby the notion of delivering a nutrient rich environment for the resident microbial communities will result in more rapid and effective elimination of contaminant compounds (Adams *et al*., 2015; Azubuike *et al*., 2016). The primary advantage of this approach is its inherent non-invasiveness. However, limitations arise from the desire to deliver bioavailable compounds into the environment and the competition generated as a result of the optimised conditions between intrinsic degraders and non-degraders (Adams *et al*., 2015). While this technique presents an environmentally friendly and cost-effective means for remediation, it has yet to show reliability and reproducibility in real world application (Macauleay & Rees, 2014; Adams *et al*., 2015).

1.7 Application of bioremediation approaches

One of the first and most notably successful field applications of bioremediation was employed in the 1989 Exxon Valdez oil spill clean-up operation in the Gulf of Alaska and Prince William Sound (Chandra *et al*., 2013). This spill involved the spillage of 42 million litres of crude oil which formed a surface slick, which occupied both marine and shoreline environments. Full-scale bioremediation of this spill primarily involved the application of two fertilizer types comprising 50,000 kg of nitrogen and 500 kg of phosphorus, to promote accelerated degradation of crude oil contaminants by naturally occurring microbes (Peterson *et al*., 2003). Bioremediation of this spill resulted in 98.7% of contaminated shoreline being designated as having no significant oil contaminants remaining within three years of the spill (Atlas, 2011). As such bioremediation of the spill was deemed successful

and garnered support and interest in bioremediation as a viable option for oil spill management (Atlas, 2011; Azubuike *et al*., 2016; Dzionek *et al*., 2016).

1.8 Bioremediation of petroleum hydrocarbons

As bioremediation is reliant on the use of organisms which display an intrinsic oil degrading capacity, a solid body of research has focused on the identification and characterisation of bacterial, yeast and fungal species which display strong potential for degrading petroleum hydrocarbons. (Das & Chandran, 2011). The degradation of these compounds by these organisms relies on their varied metabolic capabilities which are comprised of several classes of enzymes capable of facilitating mineralisation of the compounds. These capacities have been identified in yeasts, fungi and bacteria (Das & Chandran, 2011). Most studies on the use of oil degrading bacteria for petroleum bioremediation have focused on naturally occurring soil bacteria (Chaineaua *et al*., 1999; Al-Sharidah *et al*., 2000; Das & Mukherjee, 2007; Zhang *et al*., 2010; Hakima & Ian, 2017; Li *et al*., 2017). In particular, members of the mesophilic genera *Acinetobacter, Pseudomonas, Pseudonocardia, Arthrobacter, Mycobacterium* and *Bacillus* have been extensively studied (Chaineaua *et al*., 1999; Salleh *et al*., 2003; Adebusoye *et al*., 2007; Das & Mukherjee, 2007; Rojo, 2009; Das & Chandram, 2011; Muratova *et al*., 2018).

These organisms have been observed to display several strategies for enhancing the bioavailability and cellular uptake of petroleum hydrocarbons into the cell to be used as a carbon and energy source (Wentzel *et al.,* 2007). These strategies encompass the use of synthesized bioemulsifiers, cell-surface bound receptors and extracellularly secreted biosurfactants (Yakimov *et al.,* 1999; Margesin and Schinner, 2001: Wentzel *et al.,* 2007). These compounds facilitate enhanced solubility (Yakimov *et al*., 1999) direct cell- surface and indirect cell-surface interaction respectively between water insoluble petroleum hydrocarbons and the cell (Wentzel *et al.,* 2007; Kavitha *et al*., 2014).

Following cell contact with bioavailable hydrocarbons, they are transported across the plasma membrane by either passive diffusion, passive or active translocation by transporter proteins (Moreno & Rojo, 2019). Following cellular uptake, there are several mechanisms through which petroleum hydrocarbons such as n-alkanes and

isoalkanes are metabolised in bacteria (Abbasian *et al*., 2015). These include both aerobic and fermentation pathways, however aerobic pathways are more intensively studied and focused on in terms of bioremediational applications due to their broader range of substrates and reaction speeds (Rojo, 2009). Aerobic pathways primarily make use of alkane monooxygenases and cytochrome-dependant oxygenases, using O_2 to activate targeted alkane molecules (van Beilen & Funhoff, 2007; Rojo, 2009).

The most prevalent bacterial monooxygenase systems are divided into rubredoxin-dependant enzymes generally encoded by *alkB* or *alkM* genes and cytochrome P450 monooxygenases containing an alkane hydroxylase component, typically of the CYP153 family (Table 1.2) (van Beilen & Funhoff, 2005; Abbasian *et al*., 2015). Activation of long-chain alkanes by these enzymes is achieved through the cleavage of molecular oxygen and the transfer of a single oxygen atom to the target alkane. This produces a corresponding alcohol which is subjected to either terminal or sub-terminal oxidation facilitating entry into the β-oxidation pathway and incorporation into the TCA cycle (Wentzel *et al*., 2007). Less prevalent systems utilizing membrane bound (pMMO) and cytoplasmic (sMMO) methane monooxygenases have also been identified in aerobic methanotrophic bacteria (Abbasian *et al*., 2015). These facilitate activation of low chain-length alkanes by methane monooxygenases, which exhibit wide substrate ranges, primarily involved in the conversion of methane to methanol (Smith & Dalton, 2004). Novel alkane monooxygenases, LadA and AlmA, have been identified, which also facilitate long chain alkane degradation in *Geobacillus* and *Acinetobacter* spp. respectively (Feng *et al*., 2007; Wang & Shao, 2012). Alternative pathways have also been identified which utilize alkane dioxygenases, which catalyse the conversion of alkanes to corresponding hydroperoxides, prior to conversion to corresponding alcohols and aldehydes, through the Finnerty pathway (Abbasian *et al*., 2015). The predominant aerobic alkane degradation enzyme classes utilized by bacteria are identified in Table 1.2.

Table 1.2: Bacterial hydrocarbon degrading enzyme classes.

Enzyme Class	**Hydrocarbons degraded**	**Organism involved**	**Reference**
Soluble methane monooxygenases	C_{1-8} (halogenated) alkanes, alkenes and cycloalkanes	*Methylococcus, Methylosinus, Methylocystis, Methylomonas, Methylocella, Methylomirabilis*	(McDonald *et al*., 2006) (Abbasian *et al*., 2015)
Particulate methane monooxygenases	$C_{1–5}$ (halogenated) alkanes and cycloalkanes	*Methylococcus, Methylosinus, Methylocystis, Methylobacter, Methylomonas, Methylomicrobium*	(McDonald *et al*., 2006)
Alk family alkane monooxygenase	C_{5-16} alkanes, fatty acids, alkybenzenes, cycloalkanes	*Acinetobacter, Alcanivorax, Burkholderia, Geobacillus, Mycobacterium, Pseudomonas, Rhodococcus*	(van Beilen *et al*., 2003) (Abbasian *et al*., 2015)
Bacterial CYP153 cytochrome P450 monooxygenases	C_{4-16} alkanes and cycloalkanes	*Acinetobacter, Alcanivorax, Caulobacter, Mycobacterium, Rhodococcus, Sphingomonas*	(van Beilen *et al*., 2006) (Abbasian *et al*., 2015)
Dioxygenases	C_{13-44} alkanes	*Acinetobacter* sp. M-1	(Maeng *et al*., 1996)

Enzyme Class	Hydrocarbons degraded	Organism involved	Reference
LadA Terminal Long-chain alkane monooxygenase	C_{15-36} alkanes	*Geobacillus thermodenitrificans* NG80-2,	(Feng *et al*., 2007)
AlmA Terminal Long-chain alkane monooxygenase	C_{10-40} alkanes	*Acinetobacter* sp. DSM 17874, *Alcanivorax, Marinobacter*	(Thorne-Holst *et al.*, 2007; Wang & Shao, 2012)

While hydrocarbon biodegradation occurs at or near the optimal growth temperatures of these bacteria, the efficiency of hydrocarbon biodegradation has been shown to decrease with lowering temperatures (Lim *et al*., 2016). Oil viscosity increases at lower temperatures, decreasing the bioavailability of hydrocarbons (Margesin & Schinner, 2001). In addition, toxic low molecular weight hydrocarbons within petroleum mixtures present lower volatilities with decreasing temperatures, further limiting effective bioremediation (Das & Chandram, 2011). There has thus been increasing interest in the use of thermophilic bacteria for incorporation in petroleum bioremediation strategies (Vidali, 2001).

1.9 Application of thermophilic bacteria in petroleum bioremediation

Thermophilic bacteria are an increasing area of biotechnological interest due of their broad metabolic capabilities and production of thermostable enzymes (Mehta *et al*., 2016). The array of thermostable enzymes possessed by these bacteria have greater resistance not only to elevated temperatures, but also to chemical and physical denaturation (Meintanis *et al*., 2006). Substrate utilisation rates by thermophilic bacteria are typically much higher than those observed in mesophilic bacteria, indicative of faster growth rates (Mehta *et al*., 2016). Despite the potential advantages demonstrated by these bacteria and their frequent isolation from high

temperature oil-wells and subterranean oil reservoirs, limited research has been performed on thermophilic bacterial hydrocarbon biodegradation (Li *et al*., 2017).

Studies conducted on thermophilic petroleum biodegradation have focused primarily on specific thermophilic taxa residing in the genera *Anoxybacillus, Bacillus, Brevibacillus, Geobacillus, Thermoactinomyces* and *Thermus* (Meintanis *et al*., 2006; Shestakova *et al*., 2011; Elumalai *et al*., 2017). Furthermore, the focus has been on the ability of taxa in these genera to degrade specific constituents of petroleum, including monoaromatic, polycyclic aromatic and alkane hydrocarbons (Poli *et al*., 2006; Xia *et al*., 2015; Khazra *et al*., 2015: Zeinali *et al*., 2017). There is minimal data concerning the global capacity for petroleum degradation and the mechanisms concerning hydrocarbon degradation and utilisation by thermophilic taxa (Das & Chandram, 2011).

Along with enzyme pathways typically expressed in mesophilic oil hydrocarbon degrading bacteria, thermophiles often present novel or unique enzymes and pathways for degrading these compounds (Abbasian *et al*., 2015; Elumalai *et al*., 2017). For example, *Geobacillus thermodenitrificans* NG80-2, isolated from a deep-subsurface oil reservoir, has been shown to possess a novel alkane monooxygenase *ladA* gene. The LadA enzyme is involved in the degradation of $C_{15\text{-}36}$ n-alkanes, through conversion of n-alkanes to corresponding alcohols and is an example of the sparsely explored potential possessed by thermophilic bacteria (Feng *et al*., 2007; Abbasian *et al*., 2015).

1.10 Conclusions

The continual release of crude oil and petroleum derivatives into the environment poses a significant threat to maintaining a healthy and stable ecosystem. While many management and remediation strategies have been developed to respond with oil contaminated environments, most demonstrate adverse impacts on the environments undergoing treatment. Bioremediation approaches have been developed to the point where they are capable of effective innocuous remediation of oil contaminated environments. While mesophilic bioremediation application has been the focus of research, little investigation has been conducted into thermophilic approaches towards bioremediation efforts. The propensity for thermophiles to use novel, stable and efficient metabolic pathways in high temperature environments provides a yet not fully explored area for potential exploitation in bioremediation strategies. Recent identification of thermophilic alkane monooxygenase AlkB and LadA homologues in thermophilic *Geobacillus* spp. suggests the capacity for thermophilic long-chain alkane degradation capacity in thermophilic taxa. Elucidation of the capacity for oil degradation in additional thermophilic bacteria may facilitate development of thermophilic bioremediation approaches, enhancing remediation efforts.

1.11 References

Abbasian, F., Lockington, R., Mallavarapu, M. and Naidu, R., 2015. A comprehensive review of aliphatic hydrocarbon biodegradation by bacteria. Applied Biochemistry and Biotechnology, 176(3): 670-699.

Abdel-Shafy, H.I. and Mansour, M.S., 2016. A review on polycyclic aromatic hydrocarbons: source, environmental impact, effect on human health and remediation. Egyptian Journal of Petroleum, 25(1): 107-123.

Adams, G.O., Fufeyin, P.T., Okoro, S.E. and Ehinomen, I., 2015. Bioremediation, biostimulation and bioaugmentation: a review. International Journal of Environmental Bioremediation & Biodegradation, 3(1): 28-39.

Akob, D.M., Mumford, A.C., Orem, W., Engle, M.A., Klinges, J.G., Kent, D.B. and Cozzarelli, I.M., 2016. Wastewater disposal from unconventional oil and gas development degrades stream quality at a West Virginia injection facility. Environmental Science & Technology, 50(11): 5517-5525.

Atlas, R., 2011. Oil biodegradation and bioremediation: A tale of the two worst spills in U.S. history. Environmental Science and Technology, 45(16): 6709–6715.

Azubuike, C.C., Chikere, C.B. and Okpokwasili, G.C., 2016. Bioremediation techniques–classification based on site of application: principles, advantages, limitations and prospects. World Journal of Microbiology and Biotechnology, 32(11): 180.

Bacosa, H.P., Liu, Z. and Erdner, D.L., 2015. Natural sunlight shapes crude oil-degrading bacterial communities in Northern Gulf of Mexico surface waters. Frontiers in Microbiology, 6: 1325.

Baker, J.M., 2001. Oil pollution. Encyclopedia of Ocean Sciences. Shrewsbury, UK. Academic Press, 4: 1999-2007.

Beyer, J., Trannum, H.C., Bakke, T., Hodson, P.V. and Collier, T.K., 2016. Environmental effects of the Deepwater Horizon oil spill: a review. Marine Pollution Bulletin, 110(1): 28-51.

Buist, I., 2003. Window-of-opportunity for in situ burning. Spill Science & Technology Bulletin, 8(4): 341-346.

Chandra, S., Sharma, R., Singh, K. and Sharma, A., 2013. Application of bioremediation technology in the environment contaminated with petroleum hydrocarbon. Annals of Microbiology, 63(2): 417-431.

Cravo-Laureau, C. and Duran, R., 2014. Marine coastal sediments microbial hydrocarbon degradation processes: contribution of experimental ecology in the omics' era. Frontiers in Microbiology, 5: 39.

Das, N. and Chandran, P., 2011. Microbial degradation of petroleum hydrocarbon contaminants: an overview. Biotechnology Research International, 2011.

Demirbas, A. and Taylan, O., 2016. Removing of resins from crude oils. Petroleum Science and Technology, 34(8): 771-777.

Dzionek, A., Wojcieszyńska, D. and Guzik, U., 2016. Natural carriers in bioremediation: A review. Electronic Journal of Biotechnology, 19(5): 28-36.

Elumalai, P., Parthipan, P., Karthikeyan, O.P. and Rajasekar, A., 2017. Enzyme-mediated biodegradation of long-chain n-alkanes (C 32 and C 40) by thermophilic bacteria. 3 Biotech, 7(2): 116.

Enerdata, 2018. Global Energy Trends. A step backward for the energy transition?, Paris, France.

Esler, D., Ballachey, B.E., Matkin, C., Cushing, D., Kaler, R., Bodkin, J., Monson, D., Esslinger, G. and Kloecker, K., 2018. Timelines and mechanisms of wildlife population recovery following the Exxon Valdez oil spill. Deep Sea Research Part II: Topical Studies in Oceanography, 147: 36-42.

Fathepure, B.Z., 2014. Recent studies in microbial degradation of petroleum hydrocarbons in hypersaline environments. Frontiers in Microbiology, 5: 173.

Feng, L., Wang, W., Cheng, J., Ren, Y., Zhao, G., Gao, C., Tang, Y., Liu, X., Han, W., Peng, X. and Liu, R., 2007. Genome and proteome of long-chain alkane degrading *Geobacillus thermodenitrificans* NG80-2 isolated from a deep-subsurface oil reservoir. Proceedings of the National Academy of Sciences, 104(13): 5602-5607.

Ferguson, S.H., Woinarski, A.Z., Snape, I., Morris, C.E. and Revill, A.T., 2004. A field trial of in situ chemical oxidation to remediate long-term diesel contaminated Antarctic soil. Cold Regions Science and Technology, 40(2): 47-60.

Gong, Y., Zhao, X., Cai, Z., O'reilly, S.E., Hao, X. and Zhao, D., 2014. A review of oil, dispersed oil and sediment interactions in the aquatic environment: influence on the fate, transport and remediation of oil spills. Marine Pollution Bulletin, 79(2): 16-33.

Hakima, A. and Ian, S., 2017. Isolation of indigenous hydrocarbon transforming bacteria from oil contaminated soils in Libya: Selection for use as potential Inocula for soil bioremediation. International Journal of Environmental Bioremediation and Biodegradation, 5(1): 8-17.

Hesnawi, R.M. and McCartney, D.M., 2006. Impact of compost amendments and operating temperature on diesel fuel bioremediation. Journal of Environmental Engineering and Science, 5(1): 37-45.

Hoang, A.T., Pham, V.V. and Nguyen, D.N., 2018. A report of oil spill recovery technologies. Int. J. Appl. Eng. Res, 13(7): 4915-4928.

Holliger, C., Gaspard, S., Glod, G., Heijman, C., Schumacher, W., Schwarzenbach, R.P. and Vazquez, F., 1997. Contaminated environments in the subsurface and bioremediation: organic contaminants. FEMS Microbiology Reviews, 20(4): 517-523.

Ivshina, I.B., Kuyukina, M.S., Krivoruchko, A.V., Elkin, A.A., Makarov, S.O., Cunningham, C.J., Peshkur, T.A., Atlas, R.M. and Philp, J.C., 2015. Oil spill problems and sustainable response strategies through new technologies. Environmental Science: Processes & Impacts, 17(7): 1201-1219.

Mohammed, A.K., Kadhim, E.H. and Jabbar, N.M., 2018. Bioremediation of Soil Contaminated with Diesel using Biopile system. Al-Khwarizmi Engineering Journal, 14(3): 48-56.

Jernelöv, A., 2010. The threats from oil spills: now, then, and in the future. AMBIO, 39(6): 353-366.

Jiang, B., Zhou, Z., Dong, Y., Tao, W., Wang, B., Jiang, J. and Guan, X., 2015. Biodegradation of benzene, toluene, ethylbenzene, and *o*-, *m*-, and *p*-xylenes by the newly isolated bacterium *Comamonas* sp. JB. Applied Biochemistry and Biotechnology, 176(6): 1700-1708.

Ji, Y., Mao, G., Wang, Y. and Bartlam, M., 2013. Structural insights into diversity and n-alkane biodegradation mechanisms of alkane hydroxylases. Frontiers in Microbiology, 4: 58.

Jonker, M.T., Brils, J.M., Sinke, A.J., Murk, A.J. and Koelmans, A.A., 2006. Weathering and toxicity of marine sediments contaminated with oils and polycyclic aromatic hydrocarbons. Environmental Toxicology and Chemistry: An International Journal, 25(5): 1345-1353.

Kavitha, V., Mandal, A.B. and Gnanamani, A., 2014. Microbial biosurfactant mediated removal and/or solubilization of crude oil contamination from soil and aqueous phase: an approach with *Bacillus licheniformis* MTCC 5514. International Biodeterioration & Biodegradation, 94: 24-30.

Kuppusamy, S., Thavamani, P., Venkateswarlu, K., Lee, Y.B., Naidu, R. and Megharaj, M., 2017. Remediation approaches for polycyclic aromatic hydrocarbons (PAHs) contaminated soils: Technological constraints, emerging trends and future directions. Chemosphere, 168: 944-968.

Lim, M.W., Von Lau, E. and Poh, P.E., 2016. A comprehensive guide of remediation technologies for oil contaminated soil—present works and future directions. Marine Pollution Bulletin, 109(1): 14-45.

Li, X.X., Mbadinga, S.M., Liu, J.F., Zhou, L., Yang, S.Z., Gu, J.D. and Mu, B.Z., 2017. Microbiota and their affiliation with physiochemical characteristics of different subsurface petroleum reservoirs. International Biodeterioration & Biodegradation, 120: 170-185.

Macaulay, B.M. and Rees, D., 2014. Bioremediation of oil spills: a review of challenges for research advancement. Annals of Environmental Science, 8: 9-37.

Maeng, J.H., Sakai, Y., Tani, Y. and Kato, N., 1996. Isolation and characterization of a novel oxygenase that catalyzes the first step of n-alkane oxidation in *Acinetobacter* sp. strain M-1. Journal of Bacteriology, 178(13): 3695-3700.

Margesin, R. and Schinner, F., 2001. Biodegradation and bioremediation of hydrocarbons in extreme environments. Applied Microbiology and Biotechnology, 56(5-6): 650-663.

McDonald, I.R., Miguez, C.B., Rogge, G., Bourque, D., Wendlandt, K.D., Groleau, D. and Murrell, J.C., 2006. Diversity of soluble methane monooxygenase-containing methanotrophs isolated from polluted environments. FEMS Microbiology Letters, 255(2): 225-232.

Mehta, R., Singhal, P., Singh, H., Damle, D. and Sharma, A.K., 2016. Insight into thermophiles and their wide-spectrum applications. 3 Biotech, 6(1): 81.

Meintanis, C., Chalkou, K.I., Kormas, K.A. and Karagouni, A.D., 2006. Biodegradation of crude oil by thermophilic bacteria isolated from a volcano island. Biodegradation, 17(2): 3-9.

Moreno, R. and Rojo, F., 2017. Enzymes for aerobic degradation of alkanes in bacteria. Oils and Lipids; Rojo, F., Ed 1, 1-25

Moroni, D., Pieri, G. and Tampucci, M., 2019. Environmental Decision Support Systems for Monitoring Small Scale Oil Spills: Existing Solutions, Best Practices and Current Challenges. Journal of Marine Science and Engineering, 7(1): 19.

National Research Council, 2003. Oil in the sea III: inputs, fates, and effects. Washington, DC. National Academies Press.

Nyankson, E., Rodene, D. and Gupta, R.B., 2016. Advancements in crude oil spill remediation research after the Deepwater Horizon oil spill. Water, Air, & Soil Pollution, 227(1): 29.

Perelo, L.W., 2010. *In situ* and bioremediation of organic pollutants in aquatic sediments. Journal of Hazardous Materials, 177(3): 81-89.

Peterson, C., 2001. The Exxon Valdez oil spill in Alaska: acute, indirect and chronic effects on the ecosystem. Advances in Marine Biology, 39: 1-103.

Peterson, C.H., Rice, S.D., Short, J.W., Esler, D., Bodkin, J.L., Ballachey, B.E. and Irons, D.B., 2003. Long-term ecosystem response to the Exxon Valdez oil spill. Science, 302: 2082-2086.

Rehan, M., Nizami, A.S., Taylan, O., Al-Sasi, B.O. and Demirbas, A., 2016. Determination of wax content in crude oil. Petroleum Science and Technology, 34(9): 799-804.

Riazi, M.R., 2005. Characterization and properties of petroleum fractions. 1st ed. Baltimore, MD. ASTM International.

Rojo, F., 2009. Degradation of alkanes by bacteria. Environmental Microbiology, 11(10): 2477-2490.

Rubin, J., 2012. The big flatline: oil and the no-growth economy. New York, NY. St. Martin's Press.

Ruiz-Gonzalez, C., Simó, R., Sommaruga, R. and Gasol, J.M., 2013. Away from darkness: a review on the effects of solar radiation on heterotrophic bacterioplankton activity. Frontiers in Microbiology, 4: 131.

Seidel, M., Kleindienst, S., Dittmar, T., Joye, S.B. and Medeiros, P.M., 2016. Biodegradation of crude oil and dispersants in deep seawater from the Gulf of Mexico: Insights from ultra-high resolution mass spectrometry. Deep Sea Research Part II: Topical Studies in Oceanography, 129: 108-118.

Shestakova, N.M., Korshunova, A.V., Mikhailova, E.M., Sokolova, D.S., Tourova, T.P., Belyaev, S.S., Poltaraus, A.B. and Nazina, T.N., 2011. Characterization of the

aerobic hydrocarbon-oxidizing enrichments from a high-temperature petroleum reservoir by comparative analysis of DNA-and RNA-derived clone libraries. Microbiology, 80(1): 60-69.

Smith, T.J. and Dalton, H., 2004. Biocatalysis by methane monooxygenase and its implications for the petroleum industry. Studies in Surface Science and Catalysis. 151: 177-192.

Speight, J.G., 2014. The chemistry and technology of petroleum. 5th ed. Boca Raton, FL. CRC press.

Speight, J.G., 2015. Handbook of petroleum product analysis. 2nd ed. Hoboken, NJ. John Wiley & Sons.

Sun, C., Zhang, J., Ma, Q., Chen, Y. and Ju, H., 2017. Polycyclic aromatic hydrocarbons (PAHs) in water and sediment from a river basin: sediment–water partitioning, source identification and environmental health risk assessment. Environmental Geochemistry and Health, 39(1): 63-74.

Tamis, J.E., Jongbloed, R.H., Karman, C.C., Koops, W. and Murk, A.J., 2012. Rational application of chemicals in response to oil spills may reduce environmental damage. Integrated Environmental Assessment and Management, 8(2): 231-241.

Tang, J., Wang, M., Wang, F., Sun, Q. and Zhou, Q., 2011. Eco-toxicity of petroleum hydrocarbon contaminated soil. Journal of Environmental Sciences, 23(5): 845-851.

Throne-Holst, M., Wentzel, A., Ellingsen, T.E., Kotlar, H.K. and Zotchev, S.B., 2007. Identification of novel genes involved in long-chain n-alkane degradation by *Acinetobacter* sp. strain DSM 17874. Applied Environmental Microbiology, 73(10): 3327-3332.

USEPA, 2006. In situ and ex situ biodegradation technologies for remediation of contaminated sites. USEPA Engineering Issue, 62: 6-15.

Usman, M., Hanna, K. and Haderlein, S., 2016. Fenton oxidation to remediate PAHs in contaminated soils: A critical review of major limitations and counter-strategies. Science of the Total Environment, 569: 179-190.

van Beilen, J.B. and Funhoff, E.G., 2005. Expanding the alkane oxygenase toolbox: new enzymes and applications. Current Opinion in Biotechnology, 16(3): 308-314.

van Beilen, J.B. and Funhoff, E.G., 2007. Alkane hydroxylases involved in microbial alkane degradation. Applied Microbiology and Biotechnology, 74(1):13-21.

van Beilen, J.B., Funhoff, E.G., van Loon, A., Just, A., Kaysser, L., Bouza, M., Holtackers, R., Röthlisberger, M., Li, Z. and Witholt, B., 2006. Cytochrome P450 alkane hydroxylases of the CYP153 family are common in alkane-degrading eubacteria lacking integral membrane alkane hydroxylases. Applied Environmental Microbiology, 72(1): 59-65.

Van Beilen, J.B., Li, Z., Duetz, W.A., Smits, T.H.M. and Witholt, B., 2003. Diversity of alkane hydroxylase systems in the environment. Oil & Gas Science and Technology, 58(4): 427-440.

Varjani, S.J., 2017. Microbial degradation of petroleum hydrocarbons. Bioresource Technology, 223: 277-286.

Vidali, M., 2001. Bioremediation. an overview. Pure and Applied Chemistry, 73(7): 1163-1172.

Vidonish, J.E., Zygourakis, K., Masiello, C.A., Sabadell, G. and Alvarez, P.J., 2016. Thermal treatment of hydrocarbon-impacted soils: a review of technology innovation for sustainable remediation. Engineering, 2(4): 426-437.

Wang, W. and Shao, Z., 2012. Diversity of flavin-binding monooxygenase genes (*almA*) in marine bacteria capable of degradation long-chain alkanes. FEMS Microbiology Ecology, 80(3): 523-533.

Wentzel, A., Ellingsen, T.E., Kotlar, H.K., Zotchev, S.B. and Throne-Holst, M., 2007. Bacterial metabolism of long-chain n-alkanes. Applied Microbiology and Biotechnology, 76(6): 1209-1221.

Yadav, B.K., Shrestha, S.R. and Hassanizadeh, S.M., 2012. Biodegradation of toluene under seasonal and diurnal fluctuations of soil-water temperature. Water, Air, & Soil Pollution, 223(7): 3579-3588.

Yakimov, M.M., Giuliano, L., Bruni, V., Scarfi, S. and Golyshin, P.N., 1999. Characterisation of anarctic hydrocarbon-degrading bacteria capable of producing bioemulsifiers. The New Microbiologica, 22(3): 249-256.

CHAPTER TWO

SCREENING FOR THERMOPHILIC PETROLEUM HYDROCARBON DEGRADATION

2.1 Introduction

Petroleum hydrocarbon contaminated environments present unfavourable conditions which are toxic and inhospitable to most forms of life (Singh & Chandra, 2014). As such they are typically dominated by extremophiles, which are capable of enduring a wide range of adverse conditions including extremes in salinity, pH, temperature and high concentrations of toxic compounds (Margesin & Schinner, 2001). Thermophilic bacteria are frequently isolated from these types of environments, including high temperature terrestrial and marine (hypersaline) petroleum reservoirs (Margesin & Schinner, 2001; Feng *et al*., 2007; Shestakova *et al*., 2011; Guizhou *et al*., 2013). These organisms display a broad scope of survival mechanisms and possess metabolic pathways that facilitate their proliferation and tolerance of these extreme environmental conditions (Margesin & Schinner, 2001). These include thermostable metabolic pathways which allow them to mineralise and assimilate toxic compounds such as aliphatic hydrocarbons (Wentzel *et al*., 2007). Numerous studies have described the occurrence of a wide range of thermophilic bacteria, such as those in the genera *Aeribacillus, Bacillus, Clostridium, Geobacillus, Thermoleophilum, Thermomicrobium, Thermus, Thermococcus* and *Thermotoga*, in oil contaminated environments. These bacteria have further demonstrated oil degrading capacities and potential for use in bioremediation strategies (Margesin & Schinner, 2001; Feng *et al*., 2007; Wentzel *et al*., 2007; Das & Chandram, 2011; Elumalai *et al*., 2017).

The use of thermophilic bacteria in petroleum hydrocarbon remediation presents several advantages compared to mesophilic approaches. Key advantages include accelerated growth rates, more favourable enzyme kinetics and increased substrate bioavailability at higher temperatures (Wentzel *et al*., 2007). The production of biosurfactants by these thermophilic bacteria has also been observed along with their ability to degrade petroleum hydrocarbons (Mnif *et al*., 2011). Biosurfactants are a diverse group of compounds and molecules whose composition may have significant effects on the surrounding environment. These effects include alterations in surface tension, ionic interactions, solubility and emulsification properties between aqueous and non-aqueous phases (Santos *et al*., 2016). Biosurfactants enhance bioavailability of hydrocarbon compounds in aqueous

environments, improving their capacity for degradation (Manchola & Dussan, 2014). Thermophilic biosurfactant production has been observed in thermophilic strains from the genera *Aeribacillus, Bacillus* and *Geobacillus* and as such is a key area of interest in petroleum hydrocarbon degradation studies (Jara *et al*., 2013; Shekhar *et al*., 2015).

Despite the evident benefits of utilizing thermophilic bacteria for bioremediation, there have only been very limited investigations into their application in petroleum hydrocarbon bioremediation (Elumalai *et al*., 2017). In this study, a collection of >100 thermophilic aerobic bacteria were screened for their potential to degrade petroleum hydrocarbons. Our analyses found numerous strains which presented the capacity for growth in the presence of motor oil as well as its degradation in liquid-based assays. Furthermore, several species of bacteria were found to cause the displacement of plate-bound motor oil. This activity may potentially be through the production of biosurfactants, which can improve oil solubilisation and concomitantly hydrocarbon degradation. Several strains showed motor oil degradation and displacement activities of interest and will be the subject of further analysis in Chapter 3.

2.2 Methods and materials

2.2.1 Bacterial isolates

A collection of 134 thermophilic bacterial strains, isolated from a diverse range of high temperature environments, were provided by Prof. Don Cowan, Centre for Microbial Ecology and Genomics, University of Pretoria as glycerol and lyophilised stocks. Glycerol samples were re-constituted by plating on Nutrient Agar (NA) (Biolab, Merck) and Luria Bertani (LB) Agar (10.0 g/L casein hydrolysate, 5.0 g/L yeast extract, 5.0 g/L NaCl, 15.0 g/L agar bacteriological, 1.25ml 10.0% (w/v) NaOH and 995ml dH_2O). The plates were incubated at 56°C for 24-48 hours. Strains which displayed poor growth on the above agar plates were re-cultured by streaking on modified Luria Bertani Agar (mLB) (Mohr *et al*., 2018). The base medium comprises 10.0 g/L casein hydrolysate, 5.0 g/L yeast extract, 5.0 g/L NaCl, 15.0 g/L agar bacteriological, 1.25ml 10.0% (w/v) NaOH and 995ml dH_2O. After autoclaving, 1 ml of each of the following trace elements was added: 1.05M Nitrilotriacetic acid, 0.59M $MgS0_4 \bullet 7H_2O$, 0.91M $CaCl_2 \bullet 2H_2O$ and 0.04M $FeSO_4 \bullet 7H_2O$). New 25% glycerol stocks were prepared using 50% glycerol and 50% mLB broth for long-term storage. Used 15W-40 motor oil (autoclaved) was used for all tests conducted.

2.2.2 Strategy for screening of isolates for oil degrading capacities

The screening for oil degradation was conducted in three distinct phases using the assays described below.

Phase 1: The entire culture collection (134 strains) was evaluated with all assays using single replicates.

Phase 2: A sub-selection of 65 strains was screened using all assays in triplicate. Strains were selected to include strains which displayed growth on motor oil covered LB agar plates and a variety of differing morphologies between strains.

Phase 3: A further sub-selection of twenty isolates which performed optimally was evaluated again and they were identified on the basis of 16S rRNA sequencing. Strains were selected again to include a range of both apparent degraders and non-degraders. Finally, three strains were selected for further analyses (described in

Chapter 3). These strains were selected to include organisms which showed variable degradation capacity as well as close phylogenetic relatedness based on 16S rRNA sequencing results to facilitate more focused comparative genomic analyses.

2.2.3 Strain identification through 16S rRNA sequencing

The twenty strains selected for Phase 3 screening were identified by comparison of their 16S rDNA gene sequence. DNA extraction was conducted on isolated colonies using an ABIOpure Total DNA Extraction Kit (Alliance BIO, Inc., USA) according to the manufacturer's instructions, using the work-flow for Gram-positive bacteria. Extracted DNA was amplified by PCR using MyTaq™ DNA Polymerase Kit (Bioline) in 25µl reaction system containing: 5µl 5X MyTaq reaction buffer, 0,5µl universal 16S rDNA primers 8F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3'), 0,5µl MyTaq Polymerase, 16,5µl ddH_2O and 2µl DNA template. Amplification was conducted in a MyCycler Thermal Cycler (BioRad) according to the following cycles: 95°C for 5 min, 30 cycles of 95°C for 30 sec, 55°C for 30 sec and 72°C for 90 sec, followed by a final elongation at 72°C for 10 min. Sequencing of the 16S rRNA gene amplicons was done using Sanger sequencing approach at Inqaba Biotec (South Africa). Sequences obtained were compared to those on the Ribosome Database Project using the Sequence Match tool for identification (Cole *et al*., 2014). Strains were identified using the closest type strain, which were included in generation of a maximum-likelihood phylogeny. The 16S gene sequences for each strain were aligned using the Multiple Alignment using Fast Fourier Transform (MAFFT) sequence alignment server V7 (Katoh & Standley, 2013). Aligned sequences were subjected to Jmodeltest V2.1.7 to determine the most suitable evolutionary model, and a maximum-likelihood phylogeny was constructed using the PhyML web server with bootstrap support (n = 1,000 replicates) (Guindon *et al*., 2010; Darriba *et al*., 2012).

2.2.4 Oil streak assay

LB agar plates coated in 0.5ml of autoclaved used motor oil, were streak inoculated with each bacterial strain and incubated on LB agar. Streak plating with no inoculum was included as negative control. The LB agar plates were incubated at 56°C for 48 hours and evaluated for colony growth and presence of zones of

displacement on the plates. Phase 1 isolates were evaluated only once, while Phase 2 and Phase 3 strains were streaked on oil-containing LB agar plates in triplicate.

2.2.5 Oil stab assay

Oil coated LB plates were prepared as described above and inoculated with a single colony from 24 hour pre-cultures using an inoculation needle. The colony was applied to the surface of the agar at one point and incubated at 56°C for 48 hours. Plates were then evaluated for zones of oil displacement around inoculation points.

2.2.6 MacConkey bottle degradation assays

24 Hour culture assay

Screening Phase 1 and 2 did not make use of inoculum pre-cultures, instead MacConkey bottles containing 10ml of LB broth were prepared and autoclaved and subsequently inoculated with the appropriate bacterial strain. For Phase 1 and 2 screening, isolated colonies of cultures grown on LB agar plates overnight were added to the bottles and incubated at 56°C for 24 hours in a shaking incubator at 150rpm. The bottles were then spiked with 200µl of used motor oil, added to the surface of the liquid medium. These were then incubated at 56°C in a shaking incubator at 150rpm and evaluated for oil degradation activity after 24 hours and 120 hours (5 days).

For Phase 3 screening, pre-cultures were prepared in autoclaved 250ml Erlenmeyer flasks with 25ml of LB broth. These were inoculated with isolated colonies from cultures grown on LB agar plates and incubated at 56°C in a shaking incubator for 24 hours. Absorbance (growth) of the pre-cultures was determined through spectrophotometric analysis at λ_{600nm}. Volumes equivalent to 1ml of 0.5 OD_{600} of the pre-cultures were used to inoculate 10ml of autoclaved LB in MacConkey bottles. The MacConkey bottles were sealed and incubated at 56°C for 24 hours in a shaking incubator at 150rpm. These cultures were inoculated with 200µl of motor oil, incubated and evaluated for oil degradation activity at 24 hours and 120 hours post-inoculation as described above.

192 Hour culture assay

The possible effects of culture age on oil degradation capacities of strains was tested through analysis of twenty strains. Strains were cultured in MacConkey bottles in the same manner as for the 24-hour culture assays (Phase 3 screening), except for extension of the initial incubation time of 24 hours to 192 hours (8 days) prior to addition of oil.

2.3 Results and discussion

2.3.1 Strain identification through 16S rRNA gene sequencing

Sequence comparison of the 16S rRNA gene of the twenty strains selected for Phase 3 screening showed that they belong to eight distinct species in four genera, namely *Bacillus, Geobacillus, Aeribacillus* and *Anoxybacillus* (Table 2.1; Figure 2.1). Ten strains, were identified as belonging to the genus *Bacillus*. Nine of these strains, Wam9a3, But5A1A, But5A1B, Prf1A3, Wam12A2, Con1, Con2, Con4 and Con5 were identified as *Bacillus licheniformis*, displaying 95.6-99.5% identity with *B. licheniformis* DSM 13^T (Table 2.1). The remaining *Bacillus* strain (Row2A12.2) was identified as *B. smithii*, sharing 98.4% 16S rRNA sequence identity with *B. smithii* DSM 4216^T. Six strains were identified to belong to the genus *Geobacillus*, with three strains (Fur3A1B.2, Ham3A8 and DSM 465^T) belonging to *G. thermodenitrificans*. The strain G18A3 was identified as *G. vulcani,* displaying 98.4% sequence identity with *G. vulcani* 3S-1^T. *G. vulcani* has however, been reclassified as *G. thermoleovorans* (Dinsdale *et al*., 2011). G18A3 furthermore displayed 98.3% 16S rRNA sequence identity with *G. thermoleovorans* DSM 5366^T. The strain, Tok5A2 identified as *G. stearothermophilus* sharing 97.65% sequence identity with *G. stearothermophilus* DSM 22^T. The remaining *Geobacillus* strain, Tau17A2 was identified as *G. thermocatenulatus* (98.6% 16S rRNA identity with *G. thermocatenulatus* BGSC 93A1^T. Two strains, Fur6A6.1 and Wam6A1 were identified as *Aeribacillus pallidus*, displaying 96.3% and 97.6% 16S rRNA sequence identity respectively, with *A. pallidus* DSM 3670^T. The final two strains were identified as belonging to the genus *Anoxybacillus*. Ork1A6 was identified as *A. thermarum* (98.0% sequence identity with *A. thermarum* DSM 17141^T). Ork3A7 was identified as *A. mongoliensis* (97.6% 16S rRNA sequence identity with *A. mongoliensis* $T4^T$).

Table 2.1: Metadata and assay results for the twenty Phase 3 strains.

Sample	Species	Isolation Source	Streak Plate Assay			Stab Plate Assay	MacConkey Bottle 24 Hour Culture Degradation Assay		MacConkey Bottle 192 Hour Culture Degradation Assay	
			Growth	Displacement Zone	Displacement Morphology	Displacement Zone	24 Hour Degradation	120 Hour Degradation	24 Hour Degradation	120 Hour Degradation
			Phase 3	Phase 3	Phase 3	Phase 3	Phase 3	Phase 3	Phase 3	Phase 3
But5A1A	*Bacillus licheniformis*	Unknown	+++	+++	Large	+++	+++	+++	---	+++
But5A1B	*Bacillus licheniformis*	Unknown	+++	---	N/A	---	---	+++	---	+++
Con1	*Bacillus licheniformis*	Autoclaved Nigerian crude oil	+++	+++	Large	+++	+++	+++	+++	+++
Con2	*Bacillus licheniformis*	Autoclaved Nigerian crude oil	+++	+++	Large	+++	+++	+++	+++	+++
Con4	*Bacillus licheniformis*	Autoclaved Nigerian crude oil	+++	+++	Large	+++	+++	+++	---	+++
Con5	*Bacillus licheniformis*	Autoclaved Nigerian crude oil	+++	+++	Large	+++	+++	+++	+++	+++
Fur3A1B.2	*Geobacillus thermodenitrificans*	Furnas, Azores	+++	---	Large	---	+++	---	---	---
Fur6A6.1	*Aeribacillus pallidus*	Furnas, Azores	+++	+++	Small	+++	---	---	---	+++
G18A3	*Geobacillus thermoleovorans*	Sterilised Media	+++	---	N/A	---	---	+++	---	+++
Ham3A8	*Geobacillus thermodenitrificans*	Unknown	+++	---	Concentrated in/around colonies	---	---	---	---	---
Ork1A6	*Anoxybacillus thermarum*	Orakei Korako, NZ	+++	---	N/A	---	---	+++	---	---
Ork3A7	*Anoxybacillus mongoliensis*	Orakei Korako, NZ	+++	---	N/A	---	---	---	+++	+++
Prf1A3	*Bacillus licheniformis*	Praia do Fogo, Azores	+++	---	N/A	---	+++	+++	+++	+++
Row2A12.2	*Bacillus smithii*	Unknown	+++	---	N/A	---	---	---	---	+++

Sample	Species	Isolation Source	Streak Plate Assay			Stab Plate Assay	MacConkey Bottle 24 Hour Culture Degradation Assay		MacConkey Bottle 192 Hour Culture Degradation Assay	
			Growth	Displacement Zone	Displacement Morphology	Displacement Zone	24 Hour Degradation	120 Hour Degradation	24 Hour Degradation	120 Hour Degradation
			Phase 3	Phase 3	Phase 3	Phase 3	Phase 3	Phase 3	Phase 3	Phase 3
DSM 465^{T}	*Geobacillus thermodenitrificans*	Unknown	+++	---	Concentrated in/around colonies	---	---	---	---	---
Tau17 A2	*Geobacillus thermocatenulatus*	Taupo NZ	+++	---	N/A	---	---	+++	---	---
Tok5 A2	*Geobacillus stearothermophilus*	Tokaanu NZ	+++	---	N/A	---	---	+++	---	+++
Wam1 2A2	*Bacillus licheniformis*	Waimangu NZ	+++	+++	Large	+++	+++	+++	---	+++
Wam6 A1	*Aeribacillus pallidus*	Waimangu NZ	+++	---	N/A	---	---	---	---	---
Wam9 A3	*Bacillus licheniformis*	Waimangu NZ	+++	+++	Small	+++	---	+++	+++	+++

Figure 2.1: 16S rRNA phylogeny of the twenty strains used in Phase 3 screening and corresponding plate- and liquid-based assay results for each strain. Strains highlighted in yellow indicate strains subjected to screening.

2.3.2 Oil streak assay

Growth and oil displacement activity on motor oil coated LB agar plates

Phase 1 screening identified 126 bacterial strains (out of the 134 strains), which possessed the capacity to grow on the oil coated LB plates (Supplementary Table 1). Strains which did not show growth on the LB plates were not passed through successive screening assays. As oil frequently incorporates heavy metals such as lead, nickel, copper and cadmium (Vazquez-Duhalt, 1989), which are toxic to most bacteria, this analysis showed that the tested taxa had some level of tolerance to these oil constituents. While colonies could be observed on the plates inoculated with 126 of the strains, substantial differences could be observed in the distribution of the oil after incubation. In 61 (48.41%) of the 126 strains, the oil remained spread on the plate around the colonies, while for 63 other strains (50.00%), displacement of the oil away from the colonies could be observed (Figure 2.2). In a further two organisms (*G. thermodenitrificans* DSM 465^T and Ham3A8) (1.59%), the oil was observed to be concentrated in and around the growing colonies (Figure 2.3).

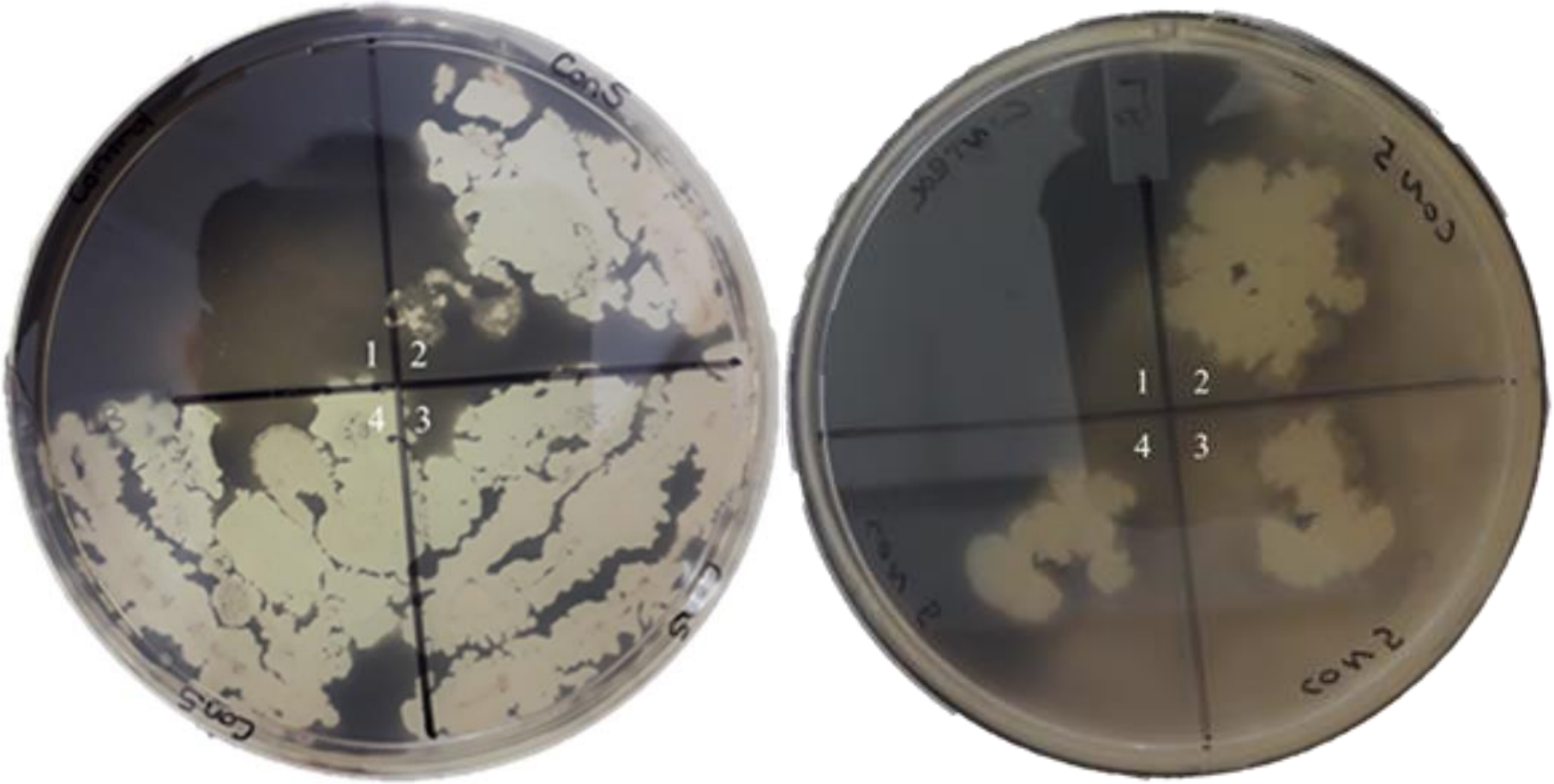


Figure 2.2: Streak plate (left) and stab plate (right) assays of *B. licheniformis* Con5 (2,3 and 4), a strain which displayed strong growth and oil displacement activity on motor oil coated LB agar plate, next to negative control (no inoculum) (1).

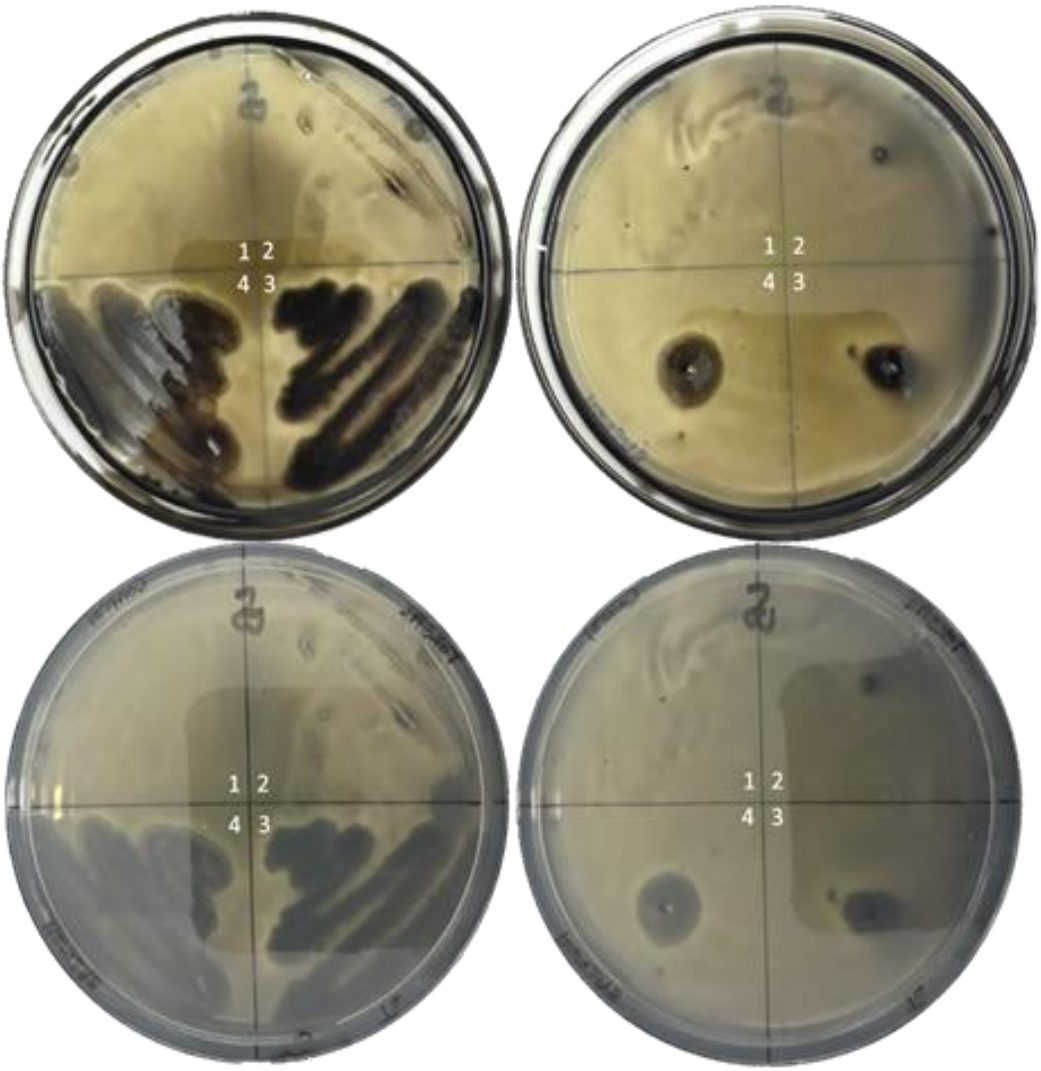


Figure 2.3: Comparative streak plate (top left and bottom left) and stab plate (top right and bottom right) for the three strains selected for further analysis in Chapter 3. Strains shown as follows: negative control (no inoculum) (1), *G. stearothermophilus* Tok5A2 (2), *G. thermodenitrificans* DSM 465^{T} (3) and *G. thermodenitrificans* Ham3A8 (4).

It must be noted that, while 50.00% of the tested strains (63 strains) showed zones of oil displacement on the oil coated LB agar plates (Supplementary Table 1), thirteen additional strains which did not cause oil displacement in the Phase 1 screening, showed oil displacement activity in the Phase 2 screening, while twenty-two displacement-positive strains from Phase 1 displayed no displacement activity in the second phase. However, thirty-three strains (50.80%) showed consistent oil

displacement activity in both Phase 1 and 2 screening. Of the twenty strains screened in Phase 3, eight strains (40.00%) displayed positive displacement results for all replicates. Six of these eight strains (75.00%) showed consistent positive displacement results in Phase 1 and Phase 3 screening.

A number of factors may be responsible for the observed oil displacement on the plates. Inconsistencies in production of zones of displacement on the motor oil plates across the different screening phases may be attributed to the use of different ages of the motor oil throughout the analyses. Oil age has been noted to affect its heavy metal composition, hydrocarbon composition and hydrocarbon chain length, affecting the polarity of oil constituents (Vazquez-Duhalt, 1989). Some bacterial taxa produce biosurfactants, which have the capacity to reduce interfacial and surface tension between polar and nonpolar molecules, facilitating the movement, and emulsification of oil in/on polar media (Manchola & Dussan, 2014). Differing hydrocarbon chain length compositions in growth media have been noted to affect biosurfactant production, which may reflect the different results obtained in terms of oil displacement due to oil ages in the different screening phases (Bordoloi & Konwar, 2009; Shekhar *et al*., 2015; Panjiar *et al*., 2017). Alternatively, oil displacement may be due to the production of extracellular enzymes or cell surface components that may change the polarity of the bacterial cell (Sepahi *et al*., 2008; Shoeb *et al*., 2012; Deliya & Jasodani, 2013; Shekhar *et al*., 2015).

2.3.3 Oil stab assay

To further validate and confirm the oil displacement capacities of the tested strains, single colonies were also stabbed into agar plates which were previously coated with motor oil.

Phase 2 screening identified twenty-five strains (38.46%) which displayed oil displacement activity around the point of inoculation (Supplementary Table 1). Nineteen of the twenty-five strains (76.00%) showed consistent positive oil displacement activity between the streak- and stab-plate assays. Of the twenty strains evaluated in Phase 3, eight showed zones of displacement around the inoculation point in the stab plate assay. These results were consistent with those

obtained with the streak plate assay, validating that these eight strains were capable of oil displacement (Supplementary Table 1).

2.4 MacConkey bottle oil degradation assays

2.4.1 24 Hour MacConkey bottle oil degradation assays

While the streak plate and stab inoculations showed the ability of the strains to grow in the presence of motor oil and validated their capacity to displace oil, it did not give a clear indication of whether the tested strains were actually capable of degrading the oil. Oil degradation capacities were thus evaluated by growing the strains in MacConkey bottles containing mLB broth and subsequently supplemented with motor oil. Phase 1 screening identified seventeen strains (13.49%) of the 126 strains tested, which displayed visible signs of oil degradation after 24 hours (Supplementary Table 1). This included the formation of particulate matter upon gentle agitation and a zone of spread oil resting on top of the aqueous medium used (Figure 2.4). Five days (120 hours) post-addition of oil, 58 strains (46.03%) of the 126 tested displayed visible signs of degradation. Phase 2 screening identified six (4.76%) of the evaluated strains which displayed visible signs of oil degradation after 24 hours. Following 120 hours of incubation seventeen strains (13.49%) showed signs of oil degradation. Phase 3 screening showed oil degradation activity in 8/20 strains (40.00%) tested after 24 hours and 13/20 strains (65.00%) for oil degradation activity after 120 hours incubation (Supplementary Table 1). As was observed for the oil displacement phenotype in the streak and stab plate assays (sections 2.3.1 and 2.3.2), some discrepancies could be observed between strains evaluated in the Phase 1, 2 and 3 screening. This may likewise be attributed to the age of the motor oil used in the screening process.

Bacterial degradation of the oils resulted in several morphologically distinct end-products. These morphologies included; a central mass with minor particulate matter in suspension, a central mass with high volumes of particulate suspension, a surface bound oil slick with particulate suspension or complete particulate suspension with no slick or central mass (Figure 2.4). The differences in morphology were likely due to a combination of the range of hydrocarbons being metabolised by the respective strains and oil displacement activity. Strains with

predicted poor or no oil degradation capacity and weak/ no oil displacement activity resulted in the formation of a central mass with large particulate matter suspended in the aqueous medium. This was observed in Phase 1 screening for 109 strains (86.50%) after 24 hours of incubation and 68 strains (53.97%) after 120 hours of incubation as well as the negative control for all phases. In Phase 2 screening, this was observed for 59 strains (90.77%) after 24 hours of incubation and 48 strains (73.85%) after 120 hours of incubation. Phase 3 screening showed seven strains (35.00%) which displayed this morphology after 24 hours incubation and six strains (30.00%) after 120 hours incubation. Those with stronger oil degradation capacities may have caused the formation of a central mass with high volumes of particulate suspension. In Phase 1 screening this was observed for eleven (8.73%) of the 126 strains tested after 24 hours and thirty-nine strains (30.95%) after 120 hours incubation. Phase 2 screening showed five strains (7.69%) which displayed this morphology after 24 hours incubation and eleven strains (16.92%) after 120 hours incubation. In Phase 3, two strains (10.00%) displayed this morphology after 24 hours incubation and four strains (20.00%) after 120 hours incubation. The presence of a surface bound oil slick with particulate suspension may indicate stronger metabolic degradation and presumed biosurfactant producing capacities by the strains producing the phenotype. Only one strain (*G. stearothermophilus* Tok5A2) displayed this phenotype after 120 hours incubation in the test phase and Phase 3 screening. Presence of complete particulate suspension with no slick or central mass was assumed to be the result of strong capacities for metabolic degradation of components of the oil. This phenotype was only seen to be produced by one strain (*B. licheniformis* But5a1B) after 120 hours incubation in screening Phase 1 and Phase 3. The visible presence of the fine particulate matter seen in these assays (Figure 2.4) may have been representative of the asphaltene component of the oil. These compounds are very stable and more resistant to degradation than the other compounds present in petroleum oils (Das & Chandram, 2011). As such, liberation of these compounds into the medium may be indicative of the strong capacity of *B. licheniformis* But5A1A for degradation of the remaining components of the motor oil. Long-chain alkane and crude oil hydrocarbon degradation activity and strong

biosurfactant production has previously been reported for *B. licheniformis* spp. and is likely the activity noted in these assays (Garcia-Alcantara *et al.*, 2016).

Depending on the quantity and qualities of potential biosurfactant production in these strains, the presence of these compounds may explain the oil morphologies observed in assays (Bordoloi & Konwar, 2009). Extracellular or membrane-bound biosurfactants may alter surface tension between the aqueous phase containing the bacterial cultures and the motor oil bodies in the MacConkey bottles. Reduction in surface tension may have caused the oil spreading activity observed. These compounds may also influence the emulsification index of the motor oil in the aqueous media, resulting in finer particulate oil drops in the media (Bordoloi & Konwar, 2009). Strains which showed potential biosurfactant production in previous assays displayed degraded oil morphologies with either no large mass or with a slick present, along with smaller particulate matter and smaller spherical oil bodies. The presence of these structures further reinforced the notion that the behaviour of the oil was, at least in part, due to biosurfactant production.

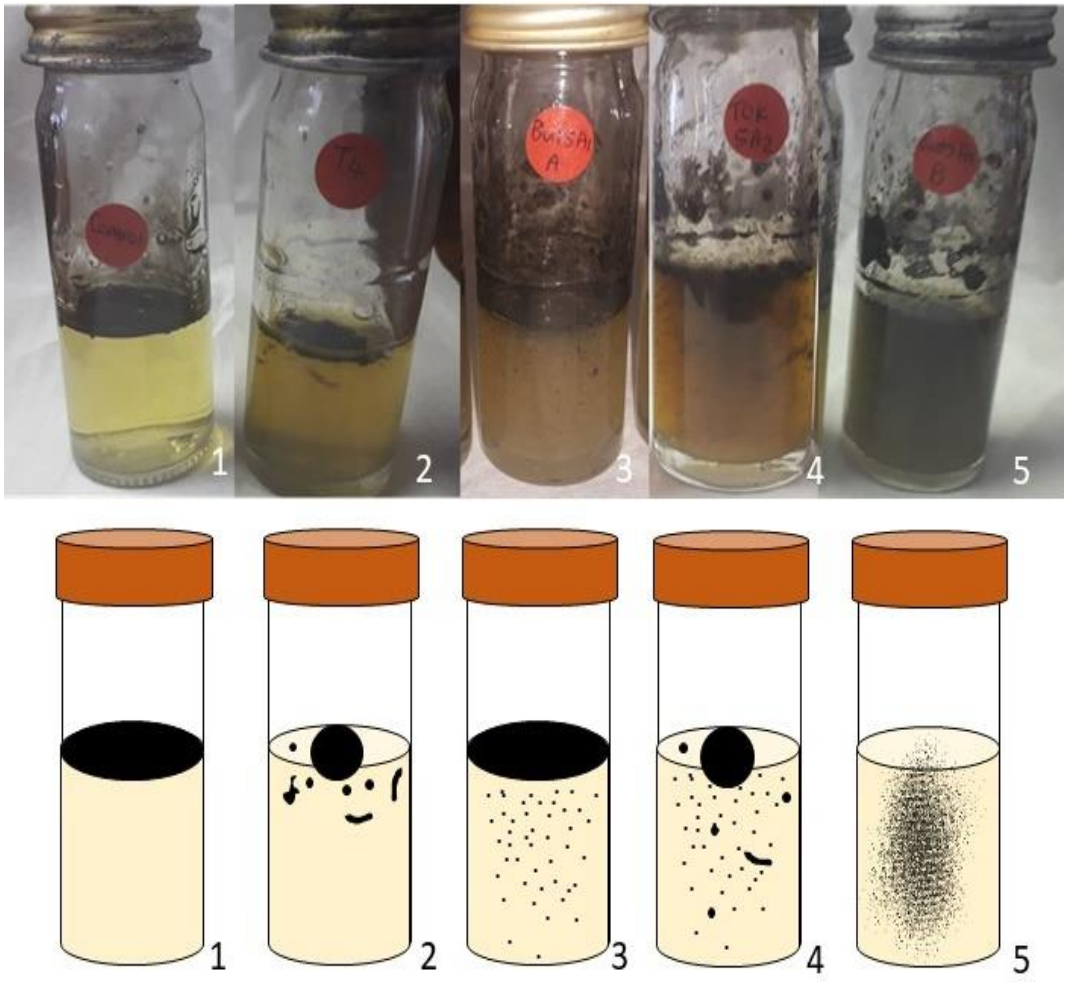


Figure 2.4: Different oil morphologies observed after 120 hour incubation with the appropriate bacterium. (Left to right) Control (1) showing no degradation; (2) *G. thermodenitrificans* DSM465^{T} showing a large mass with large particulate matter; (3) *B. licheniformis* But5a1a showing small particulate matter and spread oil slick; (4) *G. stearothermophilus* Tok5A2 showing a large oil mass with mixed particulate matter; and (5) *B. licheniformis* But5A1B showing no oil mass or slick with fine particulate matter.

2.4.2 192 Hour MacConkey bottle oil degradation assays

The effect of culture age on oil degradation capacity was evaluated using the twenty Phase 3 screening strains which were incubated for 192 hours (8 days) as compared to twenty-four hours. Eight (40%) and thirteen (65%) of the 24 hour culture assay strains exhibited oil degradation activity after 24 and 120 hours incubation with oil, respectively. Using 192-hour old cultures, six strains exhibited oil degradation activity after 24 hours and fourteen strains after 120 hours (Supplementary Table 1). For several strains culture age showed an effect on oil degradation activity. Four strains (*B. licheniformis* But5A1A, Con4 and Wam12A2, and *G. thermodenitrificans* Fur3A1B.2) showed a shift towards earlier onset of oil degradation in the younger 24 hour cultures than the 192 hour cultures. By contrast, *A. mongoliensis* Ork3A7 and *B. licheniformis* Wam9A3 showed greater oil degradation activity with the 192 hour cultures than the 24 hour cultures. Five days after oil was added, the effect of culture age was less pronounced, with five exceptions. Two strains (*A. thermarum* Ork1A6 and *G. thermocatenulatus* Tau17A2), did not exhibit oil degradation activity after 120 hours in the older cultures, however, these strains showed degradation activity when 24 hour cultures were used. Three strains, *A. pallidus* Fur6A6.1, *A. mongoliensis* Ork3A7 and *B. smithii* Row2A12.2, did not exhibit oil degradation activity after 5 days using 24 hour cultures, but showed degradation activity with the 192 hour cultures. *A. mongoliensis* Ork3A7 showed consistency in both 24 hour and 120 hour incubation periods.

Several factors may underlie the difference in oil degradation observed for the younger and older cultures. A change in metabolites present in the medium or the levels and composition of extracellular secreted compounds produced by the strains may increase bioavailability of the oil in the medium, facilitating higher degradation rates in the older cultures (Bordoloi & Konwar, 2009). Alternatively, a metabolic shift from the logarithmic growth phase to stationary phase may enable older cultures to more effectively degrade the oil constituents. Furthermore, stress experienced by the older cultures may have resulted in amplified expression of genes that increased the rate of degradation of the toxic compounds (Wentzel *et al.*, 2007).

2.6 Conclusions

Thermophilic bacteria have been shown to be ideal candidates for incorporation in oil bioremediation strategies. This can be linked in part to their frequent isolation from oil contaminated (extreme) environments where they possess the pathways for the utilisation of oil hydrocarbons as carbon source (Das & Chandram, 2011). Furthermore, the use of thermophilic microorganisms is advantageous due to increased hydrocarbon solubility and bioavailability at higher temperatures and the increased enzyme kinetics in high temperature operations (Wentzel *et al*., 2007). Here we have screened a total of 134 thermophilic bacteria isolated from a broad range of high temperature environments. A substantial number, 58 (43.28%) of the 134 evaluated strains were observed to display at least some level of degradation of motor oil. Oil degradation was highly variable among the positive strains, with some strains facilitating near-complete breakdown of the oil, with only the non-degradable asphaltenes remaining. Other strains were able to degrade some components of the oil, resulting in smaller oil flecks. A three phase screening approach identified thirteen strains that showed a consistent capacity for motor oil degradation.

The effect of culture age on oil degradation activity was assessed and was shown to produce pronounced influence in the short term, with elevated levels of degradation occurring when younger (24 hour) cultures were used for several of the strains tested. Further evaluation of the growth characteristics and metabolic states of the bacteria will need to be undertaken to better understand the reason behind this finding.

A key feature of our study was that a number of the evaluated strains caused the oil to form “balls” or emulsifications at the water-air interface, suggesting that the bacteria cause some form of encapsulation of the oil. A plausible reason for this is that these taxa synthesise an extracellular component(s), such as a biosurfactant which may influence the polarity of either the cells, medium or oil particles, facilitating encapsulation of the motor oil. Plate-based (streak and stab) assays showed that oil displacement is a prevalent feature, occurring at some level in 63 (50.00%) of the evaluated strains. The oil displacement activity of eight strains was

confirmed through three rounds of screening. This activity may furthermore reflect the oil emulsifications observed in the liquid oil degradation assays.

On the basis of the above results, three strains, all belonging to the genus *Geobacillus*, were selected for further evaluation. *G. thermodenitrificans* DSM 465^T and Ham3A8 showed poor to no capacity for oil degradation and did not display any oil displacement activity. However, both strains caused the concentration of motor oil in and around colonies when grown on motor oil coated LB agar plates. *G. stearothermophilus* Tok5A2 displayed good oil degradation activity with no oil displacement activity. In an attempt to elucidate the molecular basis between the different oil degradation capacities and displacement activities displayed by these three strains, 2D GC/MS and comparative genomic approaches were conducted in Chapter 3.

2.7 References

Bordoloi, N.K. and Konwar, B.K., 2009. Bacterial biosurfactant in enhancing solubility and metabolism of petroleum hydrocarbons. Journal of Hazardous Materials, 170(1): 495-505.

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro, A., Kuske, C.R. and Tiedje, J.M., 2013. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic acids research, 42(D1): D633-D642.

Darriba, D., Taboada, G.L., Doallo, R. and Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature methods, 9(8): 772.

Das, N. and Chandran, P., 2011. Microbial degradation of petroleum hydrocarbon contaminants: an overview. Biotechnology research international, 2011.

Dewaliya, V. and Jasodani, R., 2013. Isolation and identification of *Bacillus licheniformis* for biosurfactant production. CIBTech Journal of Microbiology, 2(4): 14-19.

Dinsdale, A.E., Halket, G., Coorevits, A., Van Landschoot, A., Busse, H.J., De Vos, P. and Logan, N.A., 2011. Emended descriptions of *Geobacillus thermoleovorans* and *Geobacillus thermocatenulatus*. International Journal of Systematic and Evolutionary Eicrobiology, 61(8): 1802-1810.

Elumalai, P., Parthipan, P., Karthikeyan, O.P. and Rajasekar, A., 2017. Enzyme-mediated biodegradation of long-chain n-alkanes (C32 and C40) by thermophilic bacteria. 3 Biotech, 7(2): 116.

Feng, L., Wang, W., Cheng, J., Ren, Y., Zhao, G., Gao, C., Tang, Y., Liu, X., Han, W., Peng, X. and Liu, R., 2007. Genome and proteome of long-chain alkane degrading *Geobacillus thermodenitrificans* NG80-2 isolated from a deep-subsurface oil reservoir. Proceedings of the National Academy of Sciences, 104(13): 5602-5607.

Garcia-Alcantara, J.A., Maqueda-Galvez, A.P., Tellez-Jurado, A., Hernandez-Martinez, R. and Lizardi-Jimenez, M.A, 2016. Maya crude-oil degradation by a

Bacillus licheniformis consortium isolated from a Mexican thermal source using a bubble column reactor. Water Air and Soil Pollution, 227(11): 413.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. and Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic biology, 59(3): 307-321.

Guizhou, G., Zheng, L., Dongfeng, Z. and Chaocheng, Z., 2013. Isolation and characterization of a thermophilic oil degrading bacterial consortium. China Petroleum Processing Petrochemical Technology, 15(2): 82-90.

Hao, R., Lu, A. and Wang, G., 2004. Crude-oil-degrading thermophilic bacterium isolated from an oil field. Canadian Journal of Microbiology, 50(3): 175-182.

Jara, A.M., Andrade, R.F. and Campos-Takaki, G.M., 2013. Physicochemical characterization of tensio-active produced by *Geobacillus stearothermophilus* isolated from petroleum-contaminated soil. Colloids and Surfaces B: Biointerfaces, 101: 315-318.

Katoh, K. and Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution, 30(4): 772-780.

Kim, I.S., Foght, J.M. and Gray, M.R., 2002. Selective transport and accumulation of alkanes by *Rhodococcus erythropolis* S+ 14He. Biotechnology and Bioengineering, 80(6): 650-659.

Manchola, L. and Dussán, J., 2014. *Lysinibacillus sphaericus* and *Geobacillus* sp biodegradation of petroleum hydrocarbons and biosurfactant production. Remediation Journal, 25(1): 85-100.

Margesin, R. and Schinner, F., 2001. Biodegradation and bioremediation of hydrocarbons in extreme environments. Applied Microbiology and Biotechnology, 56(5-6): 650-663.

Mehetre, G.T., Dastager, S.G. and Dharne, M.S., 2019. Biodegradation of mixed polycyclic aromatic hydrocarbons by pure and mixed cultures of biosurfactant

producing thermophilic and thermo-tolerant bacteria. Science of The Total Environment, 679: 52-60.

Mnif, S., Chamkha, M., Labat, M. and Sayadi, S., 2011. Simultaneous hydrocarbon biodegradation and biosurfactant production by oilfield-selected bacteria. Journal of Applied Microbiology, 111(3): 525-536.

Mohr, T., Aliyu, H., Küchlin, R., Polliack, S., Zwick, M., Neumann, A., Cowan, D. and De Maayer, P., 2018. CO-dependent hydrogen production by the facultative anaerobe *Parageobacillus thermoglucosidasius*. Microbial cell factories, 17(1): 108.

Panjiar, N., Sachan, S.G. and Sachan, A., 2017. Biosurfactants: A Multifunctional Microbial Metabolite. Microbial Applications, 2: 213-229.

Phillips, J.R., Walter, E. and Perry, J.J., 1976. *Thermomicrobium fosteri* sp. nov., a hydrocarbon-utilizing obligate thermophile. International Journal of Systematic and Evolutionary Microbiology, 26(2): 220-225.

Santos, D.K.F., Rufino, R.D., Luna, J.M., Santos, V.A. and Sarubbo, L.A., 2016. Biosurfactants: multifunctional biomolecules of the 21st century. International Journal of Molecular Sciences, 17(3): 401.

Hassanshahian, M., Emtiazi, G. and Cappello, S., 2012. Isolation and characterization of crude-oil-degrading bacteria from the Persian Gulf and the Caspian Sea. Marine Pollution Bulletin, 64(1): 7-12.

Shekhar, S., Sundaramanickam, A. and Balasubramanian, T., 2015. Biosurfactant producing microbes and their potential applications: a review. Critical Reviews in Environmental Science and Technology, 45(14): 1522-1554.

Shestakova, N.M., Korshunova, A.V., Mikhailova, E.M., Sokolova, D.S., Tourova, T.P., Belyaev, S.S., Poltaraus, A.B. and Nazina, T.N., 2011. Characterization of the aerobic hydrocarbon-oxidizing enrichments from a high-temperature petroleum reservoir by comparative analysis of DNA-and RNA-derived clone libraries. Microbiology, 80(1): 60-69.

Shoeb, E., Badar, U., Akhter, J., Ansari, F.A., Waqar, M. and Ansari, M.A., 2012. Screening of surfactant producing bacterial strains isolated from soil samples of an automobile workshop. Karachi University Journal of Science, 40(1): 31-36.

Singh, K. and Chandra, S., 2014. Treatment of petroleum hydrocarbon polluted environment through bioremediation: A review. Pakistan Journal of Biological Sciences, 17(1): 1-8.

Sorkhoh, N.A., Ibrahim, A.S., Ghannoum, M.A. and Radwan, S.S., 1993. High-temperature hydrocarbon degradation by *Bacillus stearothermophilus* from oil-polluted Kuwaiti desert. Applied Microbiology and Biotechnology, 39(1): 123-126.

Vazquez-Duhalt, R., 1989. Environmental impact of used motor oil. Science of the Total Environment, 79(1): 1-23.

Bordoloi, N.K. and Konwar, B.K., 2009. Bacterial biosurfactant in enhancing solubility and metabolism of petroleum hydrocarbons. Journal of Hazardous Materials, 170(1): 495-505.

Wentzel, A., Ellingsen, T.E., Kotlar, H.K., Zotchev, S.B. and Throne-Holst, M., 2007. Bacterial metabolism of long-chain n-alkanes. Applied Microbiology and Biotechnology, 76(6): 1209-1221.

Zarilla, K.A. and Perry, J.J., 1984. *Thermoleophilum album* gen. nov. and sp. nov., a bacterium obligate for thermophily and n-alkane substrates. Archives of Microbiology, 137(4): 286-290.

CHAPTER THREE

PHYSICOCHEMICAL AND GENOMIC CHARACTERISATION OF HYDROCARBON DEGRADATION IN THERMOPHILIC *GEOBACILLUS* ISOLATES

3.1 Introduction

Members of the thermophilic genus *Geobacillus* are common inhabitants of hydrocarbon rich environments including deep-sea oil wells, deep-subterranean oil reservoirs and high temperature oil contaminated soils (Wang *et al*., 2006; Guizhou *et al*., 2013; Hussein *et al*., 2015). Taxa in this genus are characterized by a Gram-positive cell wall structure, are rod-shaped and form spores under adverse conditions (Hussein *et al*., 2015). They are facultative anaerobes and undertake aerobic respiration and mixed acid fermentation under oxic and anoxic conditions, respectively (Shestakova *et al*., 2011; Hussein *et al*., 2015). Given their isolation from hydrocarbon contaminated environments, *Geobacillus* spp. have evolved a number of pathways that allow them to tolerate toxic hydrocarbon compounds and utilise them as carbon and energy source (Feng *et al*., 2007; Elumalai *et al*., 2017). Previous molecular studies have identified a number of genes coding for enzymes involved in hydrocarbon degradation which are present in numerous thermophiles, including *Geobacillus* spp. (Figure 3.1) (Feng *et al*., 2007; Wentzel *et al.*, 2007; Hussein *et al*., 2015). Two alkane degradation associated loci, *alkBFGHJKL* and *alkS/T* were identified and characterized in *Pseudomonas putida* Gpo1 (van Beileng *et al*., 1994). While the loci in their entirety have not been identified in *Geobacillus* spp., orthologues of the main enzyme alkane monooxygenase (AlkB) have been identified in a number of strains belonging to several distinct species, including *G. stearothermophilus*, *G. subterraneus* and *G. thermoleovorans* (Tourova *et al*., 2008). AlkB-mediated hydrocarbon degradation represents the most common aerobic pathway and has been found in a wide range of Gram-negative (e.g. *Pseudomonas, Acinetobacter, Burkholderia* spp.) as well as Gram-positive (e.g. *Mycobacterium, Nocardia* and *Prauserella*) bacteria (Wentzel *et al.*, 2007; Tourova *et al.*, 2008). Furthermore, a novel long-chain alkane degrading enzyme, LadA, was identified on the genome of *G. thermodenitrificans* NG80-2 (Feng *et al*., 2007). This novel, plasmid-encoded enzyme is involved in terminal oxidation of n-alkanes (C_{15} to C_{36}) and does not share homology with any other known alkane oxidizing enzymes (Feng *et al*., 2007; Wentzel *et al.*, 2007). A subsequent study identified and characterized three LadA-type alkane monooxygenase genes in *G. thermoleovorans* B23. These three genes were designated as *ladA*α_{B23}, *ladA*β_{B23} and

ladB$_{B23}$, and share 49.8, 34.4 and 22.7% amino acid identity with the LadA protein of *G. thermodenitrificans* NG80-2, respectively (Boonmak *et al*., 2014).

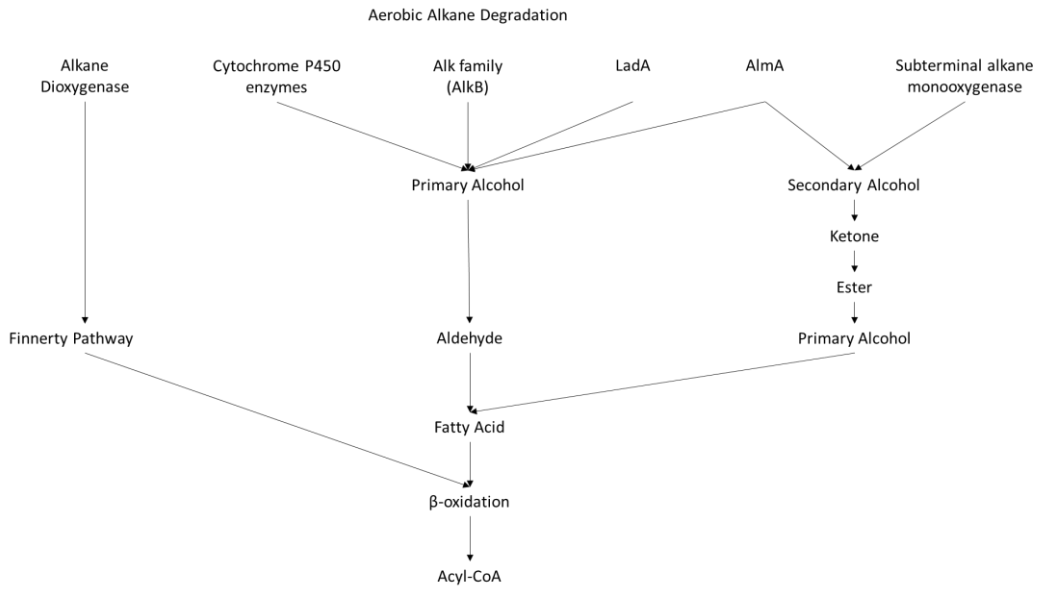


Figure 3.1: Main pathways associated with metabolism of n-alkanes in thermophilic bacteria.

In Chapter 2, a screening assay of the oil degradation capacities of an extensive set of thermophiles was undertaken. On the basis of these analyses, three isolates belonging to the genus *Geobacillus* were selected for further characterisation namely *G. thermodenitrificans* DSM 465^{T} and Ham3a8 and *G. stearothermophilus* Tok5A2. The two *G. thermodenitrificans* strains showed little capacity for oil degradation in the MacConkey bottle assays, while *G. stearothermophilus* Tok5A2 was one of the top performing strains from the assays. Selection of only *Geobacillus* strains allowed for more focused analysis of potential genetic elements and metabolic pathways that may have contributed to the variable degradative capacities displayed by the respective strains.

Here we have undertaken two-dimensional Gas Chromatography-Mass Spectrometry (2D GC/MS) to obtain qualitative and semi-quantitative data on the oil degradation capacities of the three strains. This technique allowed for the identification of a hydrocarbon chain-length profile of the oil, which was used to infer degradation of motor oil by the three strains (Hassanshahian *et al*., 2012).

Distinct patterns, in terms of hydrocarbon chain lengths, could be observed for the three strains, in line with the distinct hydrocarbon degradation capabilities observed in MacConkey bottle assays. Furthermore, we have sequenced and assembled the complete genomes of the strains Ham3a8 and Tok5A2 (the genome of *G. thermodenitrificans* DSM 465^{T} has been sequenced and is publicly available) and performed comparative genomic analyses in order to elucidate the molecular basis underlying the different oil degradation capacities of the evaluated strains.

3.2 Methods and materials

3.2.1 2D GC/MS analysis

G. thermodenitrificans DSM 465^{T} and Ham3a8 and *G. stearothermophilus* Tok5A2 were prepared, in triplicate, in MacConkey bottles as previously described in Chapter 2 (2.4.1). Negative controls containing motor oil and mLB were included in the experimental design. All samples were incubated for 10 days (240 hours) and washed twice using 10ml n-hexane (EMSURE, ACS) using 200ml separating funnels to extract the non-polar hydrocarbon components of the motor oil. After washing, the samples were placed in MacConkey bottles in a fume hood and allowed to evaporate for 48 hours to ~ 3ml volume per sample. A volume of 1.5ml of each sample was transferred to a sterile GC/MS vial for analysis. The GC/MS analysis was conducted using a Leco Pegasus 4D GCxGC-TOFMS Gas Chromatograph, using an Elite 30m x 0.25mm, 250µm column, with attached Agilent 7890 gas chromatograph autosampler. The oven temperature was set at: 80°C for 2 min, ramp 30°C/min to 240°C, ramp 3°C/min to 310 for 2 minutes (Elumalai *et al*., 2017). Helium was used as carrier gas at a constant flow rate of 1.50 ml min^{-1} and 1µl of samples were injected for analysis. All peaks except for those corresponding to alkane hydrocarbons were filtered and the area under the peaks was used to infer presence and quantity of respective length hydrocarbons.

3.2.2 Genome sequencing

DNA extraction was done using the ABIOpure Total DNA Extraction Kit (Alliance BIO, Inc., USA) using the protocol for Gram-positive bacteria. DNA was extracted in five separate reactions and pooled in a final volume of ABIOpure elution buffer. DNA extracts were evaluated for DNA concentration and quality using a Nanodrop ND1000 (Thermo Fischer Scientific, USA). Genomic DNA extractions produced good quality gDNA samples with large volumes of gDNA. Both samples showed gDNA volumes appreciably high enough for sequencing, with 260/280nm absorbance values indicative of pure quality DNA (Table 3.1).

Table 3.1: Qualities of gDNA extracts used for genome sequencing.

Strain	ng/µl	260/280	260/230
Ham3A8	393.2	1.99	1.897
Tok5A2	432.3	2.05	2.64

The genomic DNA samples were shipped for sequencing at MrDNA (Texas, USA). Libraries were prepared using Nextera DNA Flex library preparation kit (Illumina) following the manufacturer's user guide. The initial concentration of DNA was evaluated using the Qubit® dsDNA HS Assay Kit (Life Technologies). 50ng DNA was used to prepare the libraries. The samples underwent simultaneous fragmentation and addition of adapter sequences. These adapters were utilized during a limited-cycle (6 cycles) PCR in which unique indices were added to the sample. Following library preparation, the final concentration of the libraries was measured using the Qubit® dsDNA HS Assay Kit (Life Technologies), and the average library size was determined using the Agilent 2100 Bioanalyzer (Agilent Technologies). The libraries were then pooled in equimolar ratios of 1nM and sequenced for 300 cycles using the NovaSeq 6000 system with paired-end approach (Illumina).

3.2.3 Genomic assembly, annotation and comparative genomic analyses

Genome assembly and annotation

The raw genome data was first filtered to remove adapter sequences and filter out short reads (<100 bp) and those with an average Q score below 28 using the FastQ toolkit V2.2.0. Subsequently the trimmed data was assembled *de novo* using Spades V3.9.0 (Nurk *et al*., 2013). The *de novo* assembled genomes were tentatively identified using the TYpe (strain) Genome Server (TYGZ) (Meier-Kolthoff & Göker, 2019). This allowed for the identification of closely related taxa for which complete genomes are available. The genomes of the latter strains were used in reference-based assembly using Mauve V2.4.0 (Darling *et al*., 2004) and the MeDuSa genome scaffolder (Bosi *et al*., 2015). Final gap closure was performed

by BlastN analyses of scaffolded contigs against the NCBI nucleotide (nt) database (Zhang *et al*., 2000). The final genome assemblies were then submitted to the Rapid Annotations using Subsystems Technologies (RAST) server (Aziz *et al*., 2008).

Genome-based taxonomic delineation of the oil degrading *Geobacillus* strains

The high-quality draft genomes were resubmitted to the TYGZ server, which identified the strains using the Genome Blast Distance Phylogeny approach (Meier-Kolthoff *et al*., 2013; Meier-Kolthoff & Göker, 2019). Genome sequence distances were generated using the Genome Blast Distance Phylogeny pipeline. These distances were then used to infer a whole-genome based phylogeny using FastMe V2.1.6.1 (Meier-Kolthoff *et al*., 2013; Lefort *et al*., 2015; Meier-Kolthoff & Göker, 2019). Furthermore, Average Nucleotide Identity values between the sequenced strains and closest type strains were determined using OrthoANI V0.93.1 (Lee *et al*., 2015). Digital DNA-DNA Hybridization (dDDH) values were determined using the DSMZ Genome to Genome Distance Calculator (GGDC) web server, making use of formula two for incomplete genomes (Auch *et al*., 2010).

Comparative genomic analyses

The annotated protein datasets for each of the genomes generated by RAST were compared using the Sequence-based Multi Genome compare tool on the Seed Viewer server V2.0 (Overbeek *et al*., 2005). This tool performs bi-directional Blast analysis to identify orthologous and non-conserved proteins amongst the compared genomes. This means, genome elements specific to the oil degrading *G. stearothermophilus* Tok5A2 and absent from *G. thermodenitrificans* DSM 465^{T} Ham3A8 could be identified. The annotations of the genome-specific elements were examined using the RAST server (Aziz *et al*., 2008). Visual comparison of the genome sequences was done using the CGView server (Grant & Stothard, 2008). Secondary metabolites produced by the different strains were predicted on the genomes using the antiSMASH server (bacterial version) (Blin *et al*., 2019). This was done to identify any secondary metabolites encoded by the strains that possessed any known biosurfactant activities.

Identification of known oil degradation associated genes

Literature and NCBI sequence database searches identified 74 proteins involved in hydrocarbon and aromatic hydrocarbon degradation and biosurfactant production (Table 3.5). The protein sequences for each were downloaded from the NCBI server and used in local BlastP and tBlastN searches against the high-quality draft genomes with Bioedit V7.2.5 (Hall, 1999).

3.3 Results and discussion

3.3.1 2D GC/MS analysis

Three strains were selected on the basis of their distinct capacities to degrade hydrocarbons in motor oil in Chapter 2 (Figure 3.2). To gain further insights into the different hydrocarbon degrading capacities of these strains and the extent of hydrocarbon degradation, 2D GC/MS analysis was undertaken. This analysis showed several large shifts in alkane composition of the non-polar extracts.

The experimental strains showed varying alkane compositions (Figure 3.3), with several alkane chain-length peaks having decreased substantially compared to the control. Drastic decreases in peak areas were observed for C_{12}, C_{13} and C_{14} in all strains compared to the control. *G. thermodenitrificans* DSM 465^T showed the greatest decreases in these peaks with 9.91, 29.40 and 9.32% decreases in C_{12}, C_{13} and C_{14} respectively. *G. thermodenitrificans* Ham3A8 followed with 9.74, 26.38 and 8.84% decreases in C_{12}, C_{13} and C_{14} respectively. *G. stearothermophilus* Tok5A2 showed the lowest of the decreases, with 9.90, 25.36 and 9.42% decreases in C_{12}, C_{13} and C_{14} respectively. Furthermore, drastic decreases compared to the control were also observed in C_{15} and C_{16} (2.18 and 27.91% respectively) for *G. thermodenitrificans* DSM 465^T. It is notable that each strain also displayed peaks which unlike the control, showed zero value readings for their respective chain lengths. Six occurrences of this were observed with *G thermodenitrificans* DSM 465^T (C_{12}, C_{26}, C_{28}, $C_{30\text{-}31}$ and C_{43}), five with *G. thermodenitrificans* Ham3A8 for (C_{26}, $C_{30\text{-}31}$ and $C_{43\text{-}44}$) and another six with G. *stearothermophilus* Tok5A2 (C_{11}, C_{26}, C_{34}, C_{40} and $C_{43\text{-}44}$). The decreases in peak areas for the above observed alkane chain lengths provides evidence for the degradation of long-chain alkanes in these three strains. While variances were observed, none were identified to be significantly different when subjected to a student t-test using p=0,05, although were still thought to be indicative of low-level alkane degradation activity.

By contrast, numerous alkane chain-length peak areas showed increased peak areas compared to the control. The most drastic of these increases were observed for the C_{27} peaks for all three strains, which showed 74.64, 24.97 and 1.88% increases

compared to the control for *G. thermodenitrificans* DSM 465^T and Ham3A8 and *G. stearothermophilus* Tok5A2 respectively. The C_{44} peak produced by *G. thermodenitrificans* DSM 465^T was also substantially higher (8.62%) than that of the control. Only the 74.64% increase observed in C27 peak area displayed by *G. thermodenitrificans* DSM 465^T was identified to be significantly different when a student T-test was applied using p=0,05. It has previously been reported that Gram-positive bacteria, including *Bacillus* spp., are themselves capable of synthesizing high proportions of long-chain hydrocarbons (C_{11-35}), which can comprise up to 84% of their dry biomass (Ladygina *et al*., 2006). C_{27} or n-heptacosane showed the largest increases of all the peaks in each strain, particularly for the two *G. thermodenitrificans* strains. C_{27}/n-heptacosane has previously been identified as a component of a biosurfactant, lichenysin, produced by *Bacillus licheniformis* VS16 (Giri *et al*., 2017). It is thus plausible that the C_{27} peaks observed in these strains are linked to biosurfactant production, given their previously observed oil displacement activity in Chapter 2 (2.3.2). The C_{44} peak observed for *G. thermodenitrificans* DSM 465^T may likewise represent a synthesised hydrocarbon in this strain. However, synthesis of C_{44} compounds has not previously been reported in bacteria.

The reduction in hydrocarbons in the C_{12-14} range provides primary evidence for hydrocarbon degradation by all three *Geobacillus* isolates. In particular, *G. thermodenitrificans* DSM 465^T appeared to be the most promising hydrocarbon degrader, with substantially reduced peaks in the C_{12-16} range. This is in contrast to the results from the MacConkey bottle assays, where it performed the poorest. Additional peaks observed at C_{27} and C_{44} suggest production of hydrocarbons by the bacteria. It is possible that these large peaks may mask the reduction in the smaller hydrocarbon range in the remaining two strains which showed the most promising results in the MacConkey assays. The large standard errors observed were most likely the result of low replicate numbers used in conjunction with unoptimized GC/MS protocol and requires further validation of the 2D GC/MS data. To further elucidate the hydrocarbon degradation capacities of the evaluated *Geobacillus* strains, genome sequencing and comparative genomic analysis was

undertaken to characterise the molecular determinants underlying hydrocarbon degradation.

Figure 3.2: MacConkey 24 hour degradation assay results for three strains analysed in Chapter 3. Left to right: negative control, *G. thermodenitrificans* Ham3A8 and DSM 465^{T}, and *G. stearothermophilus* Tok5A2, before agitation (top) and after agitation (bottom).

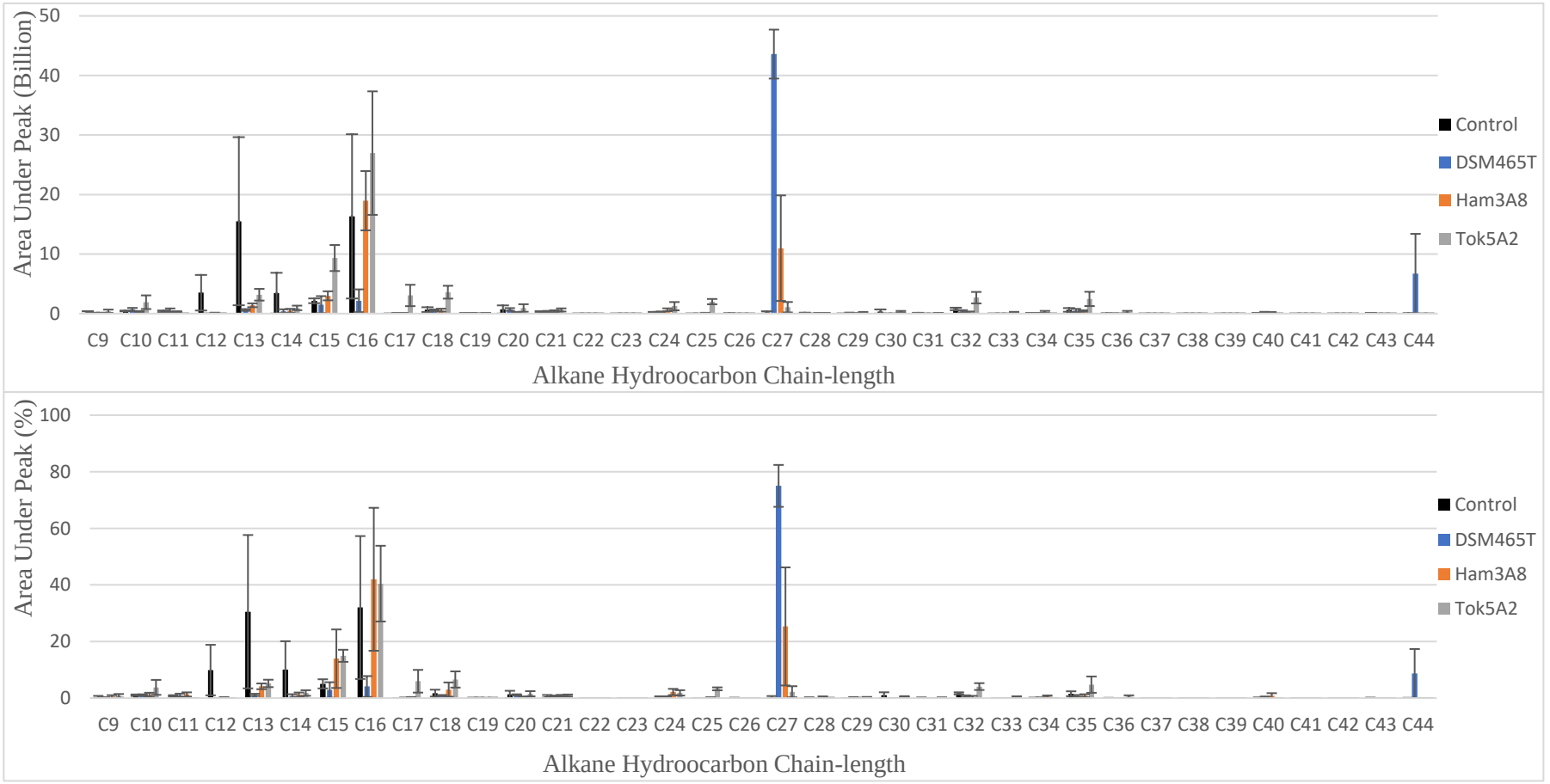


Figure 3.3: 2D GC/MS alkane carbon chain-length profile (top) and peak percentage composition of total alkane component profile (bottom), with standard errors for the control and the three experimental strains, *G. thermodenitrificans* DSM 465^{T} and Ham3A8 and *G. stearothermophilus* Tok5A2.

3.3.2 Genomic Analysis

General genome characteristics

Genome sequencing of *G. stearothermophilus* Tok5A2 and *G. thermodenitrificans* Ham3a8 yielded a total of 13,040,451 (total nucleotides 3.61 Gigabases; ~2,000X coverage) and 27,471,831 (total nucleotides 7.17 Gigabases; ~1,050X coverage) million read pairs, respectively (Table 3.2). *De novo* assembly using Spades resulted in preliminary genome assemblies comprising of 124 and 54 contigs, respectively. Reference based assembly resulted in high quality draft genomes for *G. thermodenitrificans* Ham3a8 (seven contigs) and *G. stearothermophilus* Tok5A2 (twenty-six contigs) which were used for subsequent analyses.

Table 3.2: Genome assembly metrics for *G. thermodenitrificans* Ham3A8 and *G. stearothermophilus* Tok5A2.

Strain	***G. thermodenitrificans* Ham3a8**	***G. stearothermophilus* Tok5A2**
Reads pairs (million)	27.5	13.0
Total (Gigabases)	7.17	3.62
Coverage	~2,000X	~ 1,050X
Spades contigs (>1000 bp)	54	124
N_{50} (bp)	374,157	71,408

A Genome Blast Distance Phylogeny was constructed for the strains used in this study and type strains of closely related *Geobacillus* species. The phylogeny was constructed at the whole-genome level using the Genome Blast Distance Phylogeny algorithm (Meier-Kolthoff *et al*., 2013; Meier-Kolthoff *et al*., 2019). In this phylogeny (Figure 3.4) *G. stearothermophilus* Tok5A2 clusters with *G. stearothermophilus* DSM 22^T, and *G. thermodenitrificans* Ham3A8 clusters with

G. thermodenitrificans DSM 465^{T}, suggesting they represent strains of these species. This is further supported by the phylogenomic metrics OrthoANI and dDDH. Tok5A2 shares 96.61% and 70.80% OrthoANI and dDDH with *G. stearothermophilus* DSM 22^{T}. Ham3A8 shares 99.46% and 95.60% OrthoANI and dDDH with *G. thermodenitrificans* DSM 465^{T}. Both strains displayed OrthoANI and dDDH values with their respective type strains which were above the 95-96% and 70% species thresholds for these two genome metrics, respectively (Auch *et al*., 2010; Lee *et al*., 2015).

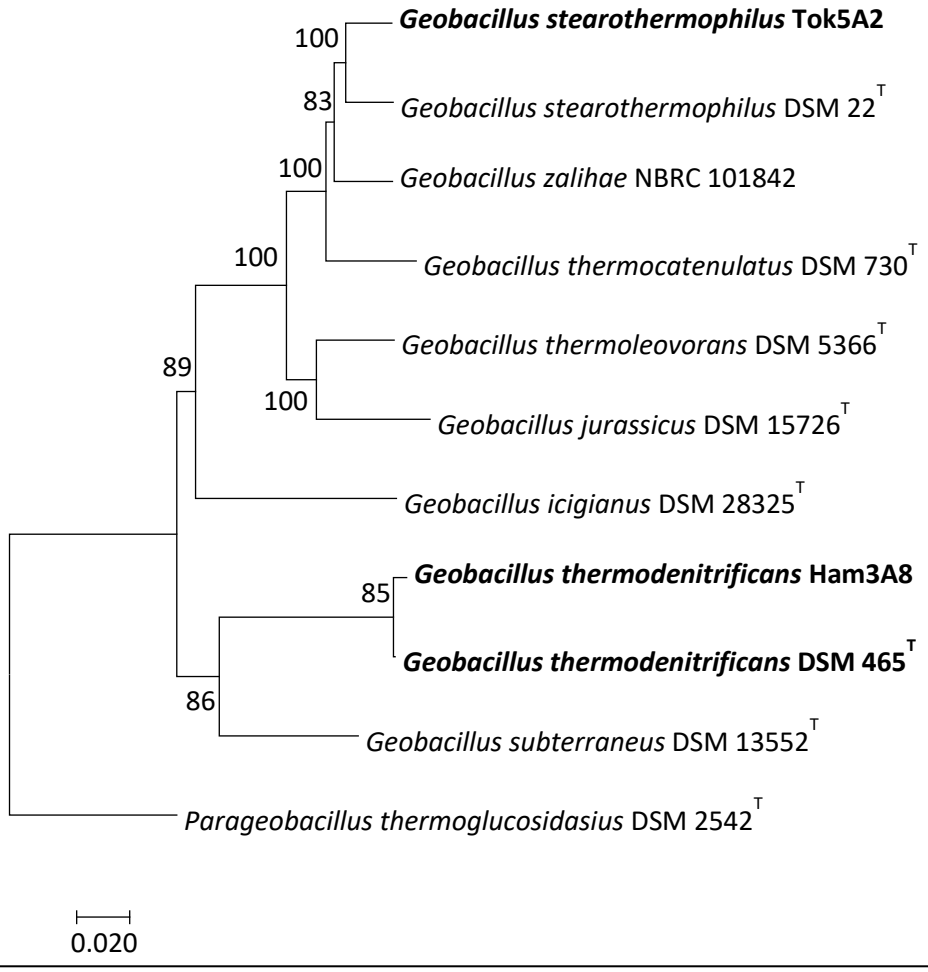


Figure 3.4: Genome Based Distance Phylogeny showing the relationships between the three final screening phase selected strains *G. thermodenitrificans* DSM 465^{T} and Ham3A8 and *G. stearothermophilus* Tok5A2 and the type strains of the genus *Geobacillus*. Strains used in subsequent analyses are in bold.

The genome of *G. thermodenitrificans* Ham3A8 is 3,611,078 bp in size with an average G+C content of 48.80% and codes for 3,984 proteins (Table 3.3). By comparison, the genome of the type strain of G. *thermodenitrificans*, DSM 465^T, is 106 kilobases smaller, codes for 172 less proteins and has a G+C content 0.3% higher than that of the former strain. The genome of *G. stearothermophilus* Tok5A2 is 71 and 176 kilobases smaller than those of *G thermodenitrificans* DSM 465^T and Ham3a8, respectively (Table 3.3). Its G+C content was substantially higher (>3% higher) than the other strains. The smaller genomes and higher G+C content were, however, typical of *G. stearothermophilus* strains (Aliyu *et al*., 2016).

Table 3.3: RAST annotation data of sequenced *Geobacillus* strains.

Strain	***G. thermodenitrificans* DSM 465^T**	***G. thermodenitrificans* Ham3a8**	***G. stearothermophilus* Tok5A2**
Genome size (bp)	3,505,276	3,611,078	3,434,618
G+C content (%)	49.1	48.8	52.3
Contig #	1	7	26
Plasmid #	-	1 (possible)	-
RNA #	118	93	108
CDS #	3,812	3,984	3,750

Comparative genomic analyses of *G. thermodenitrificans* DSM 465^T and Ham3A8 and *G. stearothermophilus* Tok5A2

The protein complements of the three *Geobacillus* strains were compared using the Seed Multi genome comparison tool (Grant & Stothard, 2008). A total of 2,997 proteins are conserved on the genomes of all three isolates, with this core set comprising between 75.23% (*G thermodenitrificans* Ham3A8) and 79.92% (*G stearothermophilus* Tok5A2) of the proteins encoded on each genome (Figure 3.5).

A substantial number of proteins (560 proteins) were also shared by the two *G. thermodenitrificans* strains but not *G. stearothermophilus* Tok5A2. When considering the unique fraction of each genome, 6.06, 9.54 and 18.19% of proteins were unique to *G. thermodenitrificans* DSM 456^{T}, Ham3A8 and *G. stearothermophilus* Tok5A2, respectively (Figure 3.5). A visual comparison of the genomes using the CGView server (Grant & Stothard, 2008), shows large non-conserved regions occurring on the genomes, which may reflect genomic islands that have led to genome diversification (Figure 3.6).

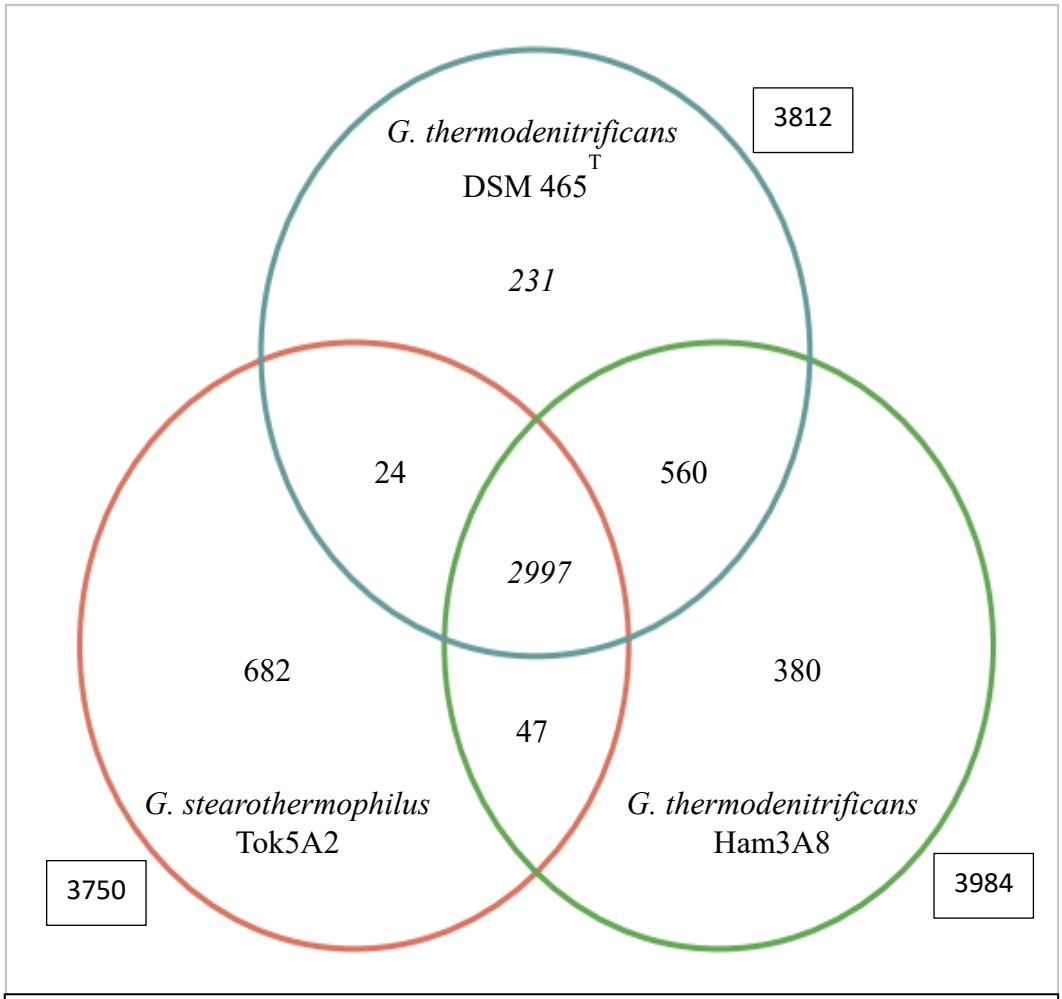


Figure 3.5: Venn diagram showing coding domain sequences shared and exclusive to *G. thermodenitrificans* DSM 465^{T} and Ham3A8 and *G. stearothermophilus* Tok5A2.

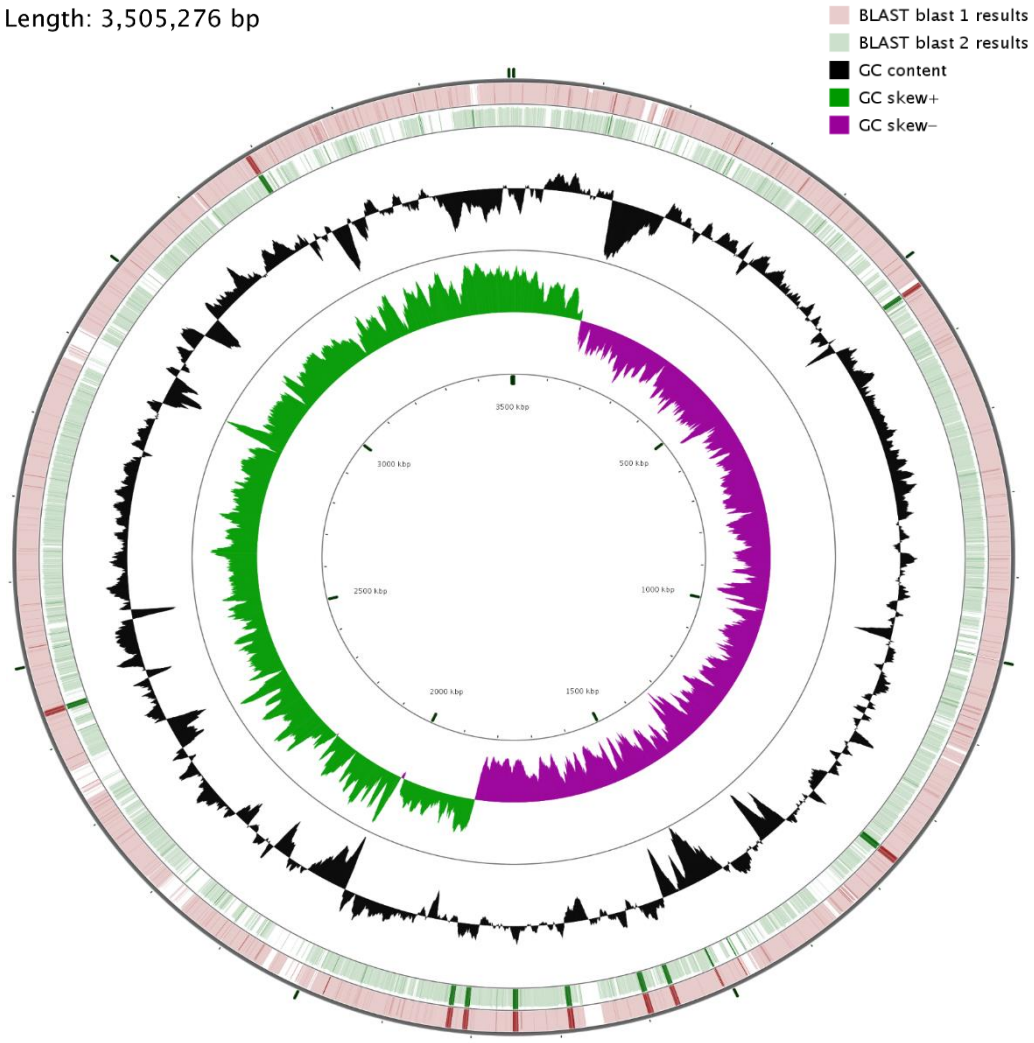


Figure 3.6: Diagram showing core genome G+C distribution comparing *G. thermodenitrificans* DSM 465^T, *G. thermodenitrificans* Ham3A8 (Blast 1: red outer ring) and *G. stearothermophilus* Tok5A2 (Blast 2: green inner ring), using *G. thermodenitrificans* DSM 465^T as the reference genome. Conserved protein coding sequences are seen as red/green respectively while non-conserved sequences are shown as white.

Comparison of the genomes on the basis of the RAST subsystems, showed that the *G. stearothermophilus* Tok5A2 encodes substantially fewer proteins involved in carbohydrate metabolism (29 and 27 less), amino acid metabolism (29 and 30 less), aromatic compound metabolism (7 and 15 less), membrane transport (19 less), cell wall and capsule biosynthesis (6 and 16 less) than *G. thermodenitrificans* DSM 465^T and Ham3A8, respectively. Furthermore, no proteins involved in iron

acquisition and metabolism were identified in the genome of the former strain, unlike *G. thermodenitrificans* DSM 465^T and Ham3A8, which code for 24 and 13 proteins involved in this function, respectively. However, upon further inspection nine genes coding for heme biosynthetic protein orthologues were identified on the genomes of all three strains (Dailey *et al.*, 2017). As such it is likely that each strain is capable of heme synthesis. The genome of *G. stearothermophilus* Tok5A2 codes for substantially higher numbers of proteins involved in cofactor, vitamin, prosthetic group and pigment synthesis (4 and 17 more), virulence, disease and defence (16 and 15 more), protein metabolism (3 and 15 more) and DNA metabolism (9 and 8 more) than *G. thermodenitrificans* DSM 465^T Ham3A8 strains respectively. By contrast, limited differences in the distribution of protein functional categories could be observed between the two *G. thermodenitrificans* strains.

Putative pathways of alkane degradation from the annotated *Geobacillus* genomes

The RAST annotation pipeline does not incorporate pathways associated with alkane degradation other than elements of the alkane β-oxidation pathway involved in secondary metabolism of Acyl CoA (Wentzel *et al.*, 2007). Among the subsystems identified by RAST were those involved in the degradation of aromatic compounds and alkanesulfonates. Enzymes involved in the central degradation pathways of gentisate and salicylate, peripheral degradation of quinate and of sulphur containing alkanesulfonates were found to be encoded on the genomes of all three strains.

β-oxidation involves the breakdown of fatty acids produced by pathways such as alkane degradation, where fatty acyl CoA is metabolised and its products fed into the Krebs cycle (Abbasian *et al*., 2015). Key proteins involved in this pathway are long chain fatty acid CoA ligases, enoyl CoA hydratases, 3-hydroxyacyl CoA dehydrogenases and 3-ketoacyl CoA thiolases. These proteins are involved in the successive stepwise metabolism of fatty acyl CoA (Jimenez-Diaz *et al.*, 2019). A full complement of genes encoding protein orthologues of the β-oxidation pathway were identified on the genomes of *G. thermodenitrificans* DSM465^T and Ham3A8

and *G. stearothermophilus* Tok5A2. Thirteen orthologues of these proteins were encoded on the genome of *G. thermodenitrificans* Ham3A8 and *G. stearothermophilus* Tok5A2 while the genome of *G. thermodenitrificans* DSM465^T encoded ten orthologues. Three long chain fatty acid CoA ligases were identified on the genome of *G. thermodenitrificans* DSM465^T (DSM465_CDS663/CDS667/CDS2842) and four on the genomes of *G. thermodenitrificans* Ham3A8 (Ham3A8_CDS910/CDS2404/CDS2585/CDS2588) and *G. stearothermophilus* Tok5A2 (Tok5A2_CDS800/CDS905/CDS2325/ CDS2328), respectively. Three enoyl CoA hydratases were identified on the genome of *G. thermodenitrificans* DSM465^T (DSM465_CDS2844/CDS3617-3618) and Ham3A8 (Ham3A8_CDS912/CDS1711-1712), respectively and four on the genome of *G. stearothermophilus* Tok5A2 (Tok5A2_CDS803/CDS907 /CDS1580-1581). Two 3-hydroxyacyl CoA dehydrogenases were identified on the genomes of *G. thermodenitrificans* DSM465^T (DSM465_CDS662/CDS3620) and *G. stearothermophilus* Tok5A2 (Tok5A2_CDS1583/CDS2324), respectively and three on the genome of *G. thermodenitrificans* Ham3A8 (Ham3A8_CDS1714/ CDS2408/CDS2584). Three 3-ketoacyl CoA thiolases were identified on the genome of *G. thermodenitrificans* DSM465^T (DSM465_CDS282/CDS660/ CDS3621) and four on the genomes of *G. thermodenitrificans* Ham3A8 (Ham3A8_CDS1715/CDS2182/CDS2403/CDS2583) and *G. stearothermophilus* Tok5A2 (Tok5A2_CDS805/CDS1584/CDS1969/CDS2323). As such it is likely that all three strains are capable of facilitating β-oxidation of alkane degradation pathway products.

Gentisate and salicylate have been identified as intermediates of polyaromatic hydrocarbon (PAH) and naphthalene degradation, with salicylate also serving as an up-regulator of PAH degrading activity in *Pseudomonas* spp. (Grund *et al*., 1992; Chen & Aitken, 1999). Two salicylate degradation-associated fumarylacetoacetate hydrolase family proteins were identified on the genomes of each strain (DSM465^T_CDS1254/CDS3676; Ham3A8_CDS1778/CDS3189; Tok5A2_CDS2110/CDS2972). *G. thermodenitrificans* Ham3a8 uniquely codes for two gentisate 1,2-dioxygenases (EC 1.13.11.4; Ham3A8_CDS2389 and Ham3A8_CDS2401). Gentisate 1,2-dioxygenases catalyse the oxygen dependent

ring-cleavage of gentisate, breaking its ring structure, facilitating further degradation of gentisate to pyruvate. (Eppinger & Stolz, 2017). The presence of orthologues of these enzymes suggest that these strains may be able to utilize gentisate and salicylate, known intermediates of petroleum oil PAH degradation, as a carbon source (Grund & Kutzner, 1999; Chen & Aitken, 1999; Phale *et al*., 2007).

Quinate is a ubiquitous monoaromatic hydrocarbon, produced by plants and bacteria. A protein coding for an orthologue of a 3-dehydroquinate dehydratase (EC 4.2.1.10) was identified in each *Geobacillus* strain (DSM465^T_CDS3602; Ham3A8_CDS1691; Tok5A2_CDS1561). This enzyme is involved in the interconversion of 3-dehydroquinic acid to 3-dehydroshikimic acid, an intermediate of aromatic amino acid synthesis and formation of protocatechuic acid which is fed into the central carbon metabolism pathways (Grund & Kutzner, 1999).

The RAST subsystems identified eight, nine and eight proteins in the genomes of *G. thermodenitrificans* DSM 465^T and Ham3A8 and *G. stearothermophilus* Tok5A2, respectively involved in alkanesulfonate degradation. Alkanesulfonates are organosulfur compounds found in petroleum rich environments, where they are degraded as a source of sulphur by a range of bacteria, including thermophilic bacteria (Boonmak *et al*., 2014; Zheng *et al*., 2016; Peng *et al*., 2019). The genomes of all three strains code for at least one (DSM465^T_CDS2856 and Ham3a8CDS_922) SsuD alkanesulfonate monooxygenase (EC 1.14.14.5) with *G. stearothermophilus* Tok5A2 coding for two (Tok5A2_CDS336 and Tok5A2_CDS915). SsuD catalyses the conversion of $C_{1\text{-}14}$ alkanesulfonates to corresponding aldehydes and sulphites in conjunction with NAD(P)H:flavin oxidoreductase SsuE (EC 1.5.1.29) (Eichhorn *et al*., 1999; Armacost *et al*., 2014). The genome of only *G. thermodenitrificans* Ham3A8 codes for a SsuE orthologue (Ham3A8_CDS305). The genomes of the two *G. thermodenitrificans* strains also encode an alkyl hydroperoxide reductase subunit C-like protein (AhpC-like) (DSM465^T_CDS937 and Ham3A8_CDS2857) which was not present on the genome of *G. stearothermophilus* Tok5A2. AhpC proteins are involved in the reduction of hydrogen peroxides (Seaver & Imlay, 2001). Upon further analysis of the coding sequences co-localised with the above genes, coding sequences which corresponded to putative *Geobacillus* spp. SsuCBAED protein orthologues were

identified. These were present on the genomes of *G. thermodenitrificans* DSM 465^T (DSM465^T_CDS2850-2854) and Ham3A8 (Ham3A8_CDS918-922). *G. stearothermophilus* Tok5A2 however, only possessed orthologues of SsuA, SsuD and SsuE (Tok5A2_CDS911, Tok5A2_CDS 912 and Tok5A2_CDS 915). The *ssuEADBF* and *ssuEADCB* operons from *Pseudomonas putida* S-313 and *Escherichia coli* EC1250 have been identified to code for an ATP-binding cassette-type transporter (SsuABC) and a two-component reduced flavin mononucleotide-dependent monooxygenase (SsuED) involved in alkanesulfonate degradation (van der Ploeg *et al*., 1999; Kahnert, *et al*., 2000). The presence of orthologous proteins for alkanesulfonate degradation suggests the *Geobacillus* spp. have evolved the capacity to use these compounds as a carbon source, a pertinent capacity for growth in sulphur rich petroleum oil environments (Speight, 2014).

Toluene, an alkyl benzene compound and benzoate, a very basic monoaromatic compound, are targeted problematic compounds commonly associated with petroleum oil spills (Jaggi *et al*., 2017). A toluenesulfonate zinc-independent alcohol dehydrogenase was identified on the genome of both *G. thermodenitrificans* DSM 465^T (DSM465^T_CDS1620) and Ham3A8 (Ham3A8_CDS3612). This enzyme catalyses the conversion of a sulfobenzyl alcohol to a corresponding sulfobenzyl aldehyde and falls under the pathway converting toluene sulfonate to benzoate (Cook *et al*., 1999). Two orthologues of the Major Facilitator Superfamily (MFS) transporter protein BenK, which is involved in the specific transport of benzoate into the cytoplasm (Pao *et al*., 1998), are encoded on the genome of *G. thermodenitrificans* Ham3A8 (Ham3A8_CDS301 and Ham3A8_CDS2406).

Genome identification of secondary metabolites with putative roles in hydrocarbon degradation

Bacterially synthesised secondary metabolites have a broad range of functions providing them with resistance, antagonistic and metabolic pathways that provide them with selective advantages over other microorganisms in the environment (Ziemert *et al*., 2016). Secondary metabolites such as biosurfactants, solvents and biopolymers have been observed to increase the effectivity of bioremediation efforts of petroleum oil hydrocarbons (Varjani & Uspani, 2017). The genomes of

the three *Geobacillus* strains relevant to this study were analysed for secondary metabolites using the antiSMASH pipeline (Blin *et al*., 2019). This approach identified five, four and five secondary metabolite clusters in *G. thermodenitrificans* DSM 465^{T} and Ham3A8 and *G. stearothermophilus* Tok5A2, respectively (Table 3.4). The two *G. thermodenitrificans* strains incorporate a gene cluster coding for the synthesis of the lanthipeptide Geobacillin I. This lanthipeptide possesses antimicrobial activity, inducing bacterial membrane pore formation when complexed with a lipid II (Garg *et al*., 2014). While *G. stearothermophilus* Tok5A2 appeared to be missing this gene cluster, it coded for two bacteriocins involved in bacterial antagonism. The first predicted bacteriocin shares orthology with the antilisterial Linocin M18 of *Brevibacterium linens* M18 (Valdes-Stauber & Scherer, 1996). The second predicted bacteriocin is an orthologue of Lactococcin 972 of *Lactococcus lactis* subsp. *lactis* IPLA 972. Lactococcin 972 is a non-lantibiotic bacteriocin with abnormal mode of action, where instead of forming pores in sensitive bacterial strain cytoplasmic membranes, inhibits septum synthesis, preventing cell division (Martinez *et al*., 2008). *G. thermodenitrificans* DSM 465^{T} possesses a 49,724 nt gene cluster coding for a predicted non-ribosomal peptide siderophore which shares 56% similarity with the bacillibactin siderophore cluster of *Bacillus subtilis* subsp. *subtilis* 168. This metabolite is produced under iron starvation conditions and facilitates intracellular localization of chelated iron by ABC transporters (Miethke *et al*., 2006). All three of the evaluated strains also incorporate a gene cluster with a predicted role in the synthesis of a hopanoid terpene. Hopanoid biosynthesis in bacteria is associated with elevated antibiotic resistance and stress tolerance (Saenz *et al*., 2015). While the secondary metabolite features above contribute to provide competitive advantages to the strains that encode them, none of them are predicted to play a role in hydrocarbon degradation.

Table 3.4: Secondary metabolite gene clusters identified in *G. thermodenitrificans* DSM 465^T and Ham3A8 and *G. stearothermophilus* Tok5A2.

Metabolite	Organism	Locus Tag	Similar Metabolite Cluster	Similarity (%)	Reference
Betalactone	*G. thermodenitrificans* DSM 465^T	Chrom1: 303875-328179	Fengycin	46	(Koumoutsi *et al.*, 2004)
	G. thermodenitrificans Ham3A8	Chrom1C: 532111-556415	Fengycin	46	
	G. stearothermophilus Tok5A2	Chrom1J: 590-24792	Fengycin	46	
T3PKS	*G. thermodenitrificans* DSM 465^T	Chrom1: 333981-375048	-	-	-
	G. thermodenitrificans Ham3A8	Chrom1C: 5622127-603284	-	-	-
	G. stearothermophilus Tok5A2	Chrom1J: 32026-73144	-	-	-
Lanthipeptide	*G. thermodenitrificans* DSM 465^T	Chrom1: 1509744-1536156	Geobacillin 1	100	(Garg *et al.*, 2012)
	G. thermodenitrificans Ham3A8	Chrom1C: 1804095-1830507	Geobacillin 1	100	
Non-Ribosomal Peptide Synthetase	*G. thermodenitrificans* DSM 465^T	Chrom1: 1638735-1688458	Bacillibactin	53	
Terpene	*G. thermodenitrificans* DSM 465^T	Chrom1: 3324927-3346780	-	-	-
	G. thermodenitrificans Ham3A8	Chrom1C: 47376-69229	-	-	-

Metabolite	Organism	Locus Tag	Similar Metabolite Cluster	Similarity (%)	Reference
	G. stearothermophilus Tok5A2	Chrom1F: 810192-831787	-	-	-
Bacteriocin	*G. stearothermophilus* Tok5A2	Chrom1L: 93158-104012	-	-	-
		Chrom1M: 20892-31263	-	-	-

However, the genomes of all three strains also incorporate a betalactone cluster, which is predicted to share 46% gene cluster similarity with the fengycin biosynthetic cluster (*fenABCDE*) in *Bacillus velezensis* FZB42 (Figure 3.7). Fengycin is an antifungal lipopeptide which has been noted to display biosurfactant activity. However, on closer inspection, this cluster similarity was linked to flanking genes of the fengycin cluster, *yngEFGHIJ*, which are predicted to be involved in leucine degradation during sporulation.

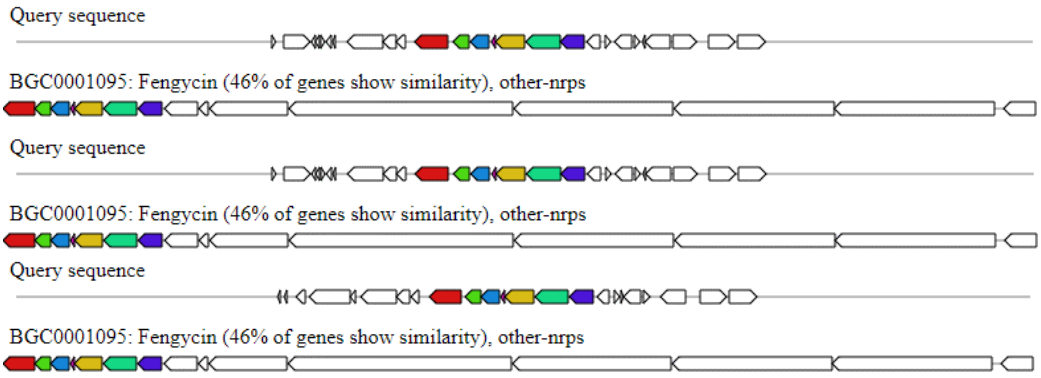


Figure 3.7: Diagram showing proteins identified to show similarity to Fengycin gene cluster in *Bacillus velezensis* FZB42. (Top) *G. thermodenitrificans* DSM 465^{T}. (Middle) *G. thermodenitrificans* Ham3A8. (Bottom) *G. stearothermophilus* Tok5A2.

In addition to the secondary metabolites identified using the antiSMASH pipeline, a further predicted secondary metabolite could be identified from the RAST annotated genome of all three strains. The RAST annotation identified orthologues of the proteins SpaB, SpaC and SpaT in the genomes of each *Geobacillus* strain (DSM465^T_CDS1647-1649, Ham3A8_CDS3638-3641 and Tok5a2_CDS3358-3362). The *spaBTC* translational unit in *B. subtilis* subsp. *spizizenii* ATCC 6633 has been associated with the posttranslational modification and transport of subtilin (Stein *et al*., 2003). Subtilin is a class I lanthionine antibiotic produce by *B. subtilis* subsp. *spizizenii* ATCC 6633, which induces pore formation in the cytoplasmic membrane of sensitive Gram positive bacteria (Klein & Entian, 1994; Stein *et al*., 2003). Along with antibiotic activity, subtilin has also been observed to possess strong biosurfactant activities as a result of its strong surface net hydrophobicity (Kuboi *et al*., 1994).

The subtilin gene cluster in *B. subtilis* subsp. *spizizenii* ATCC 6633 has been identified as four transcriptional units, *spaBTC*, *spaS*, *spaIFEG* and *spaRK*. The four translational units code for posttranslational modification and transport proteins, the subtilin prepeptide encoding gene, immunity proteins and a two-component regulatory system respectively (Stein *et al*., 2003). Upon further examination of the surrounding RAST annotated CDS's, along with SpaBTC orthologues (DSM465^T_CDS1647-1649, Ham3A8_CDS3638-3639 and Ham3A8_CDS3641 and Tok5A2_CDS3358, Tok5A2_CDS3360 and Tok5A2_CDS3362) further subtilin cluster genes were identified. These included genes encoding six proteins orthologous to the SpaSFEGRK proteins of *B. subtilis* subsp. *spizizenii* ATCC 6633 (AAB91589, AAB91597, AAA22774, AAB91595, AAB91594 and AAB91593). The SpaSFEGRK orthologues coded by *G. thermodenitrificans* DSM 465^T (DSM465^T_CDS1650/1641/1649/1643/1646/1645), *G. thermodenitrificans* Ham3A8 (Ham3A8_CDS3642/3632/3641/3634/3637/3636) and *G. stearothermophilus* Tok5A2 (Tok5A2_CDS3378/3352/3362/3354/3357) share 43.78, 43.89 and 43.60% average amino acid identity to SpaSFEGRK of *B. subtilis* subsp. *spizizenii* ATCC 6633. The presence of orthologues of subtilin biosynthetic proteins suggests that the *Geobacillus* strains in this study may produce an

analogous secondary metabolite, which may play a potential role in hydrocarbon degradation and the oil displacement phenotypes observed in the plate- and liquid-based assays in Chapter 2 (2.3.2 and 2.4.1).

Genome screening for known hydrocarbon degradation enzymes

In addition to the analyses of the RAST annotated genomes above, the genomes of the three *Geobacillus* strains in this study were evaluated for the presence or absence of 74 proteins with known function in hydrocarbon degradation (Table 3.5). Of the 74 proteins evaluated, orthologues of one, five and seven proteins were identified in *G. thermodenitrificans* DSM 465^T, Ham3A8 and *G. stearothermophilus* Tok5A2, respectively. No biosurfactant production related protein orthologues were identified.

Genes associated with n-alkane degradation. Orthologues of six long-chain alkane degradation associated proteins were identified on the genome of *G. stearothermophilus* Tok5A2. Three orthologues, sharing 100% amino acid identities, to $LadA\alpha_{B23}$, $LadA\beta_{B23}$ and $LadB_{B23}$ of *G. thermoleovorans* B23 (BAM76377, BAM76372 and BAM76371) were encoded on the genome of *G. stearothermophilus* Tok5A2 (Tok5A2_CDS330, Tok5A2_CDS335 and Tok5A2_CDS336). $LadA\alpha_{B23}$ and $LadA\beta_{B23}$ are long-chain alkane monooxygenase enzymes, identified and functionally characterised as LadA homologues (Boonmak *et al*., 2014). LadA is a long chain alkane monooxygenase responsible for the conversion of C_{15-36} alkanes to their corresponding primary alcohols in the LadA alkane degradation pathway in *G. thermodenitrificans* NG80-2 (Feng *et al*., 2007). $LadB_{B23}$ is an alkane monooxygenase representative of a domain group distinct from the highly similar LadA alkane monooxygenase and SsuD alkanesulfonate monooxygenase domain groups (Boonmak *et al*., 2014).

Another 496aa protein (Tok5A2_CDS810) encoded on the genome of *G. stearothermophilus* Tok5A2 shared a 99.40% amino acid identity (Bitscore: 979; E-value: 0.0) with that of an aldehyde dehydrogenase (Bt-Aldh) of *G. thermoleovorans* B23 (BAB16600). This enzyme is involved in the oxidation of aldehydes (preferentially long-chain aldehydes) to their corresponding carboxylic acids (Kato *et al*., 2010). In the latter organism Bt-Aldh is upregulated during long-

chain alkane degradation commencement by *G. thermoleovorans* B23 and its upstream location to other alkane inducible protein P21 suggest its involvement in a long-chain alkane degradation pathway (Kato *et al*., 2010). The genome of *G. stearothermophilus* Tok5A2 also encodes two proteins (Tok5A2_CDS526 and Tok5A2_CDS812) which share 91.33% (Bitscore: 350; E-value: 4.00E-098) and 99.50% (Bitscore: 387; E-value: 2.00E-109) amino acid identity with P16 and P21 membrane proteins of *G. thermoleovorans* B23 (BAB16679 and BAB16602), respectively. P16 and P21 are membrane proteins which are upregulated upon commencement of alkane degradation by *G. thermoleovorans* B23 (Kato, et al., 2009). These proteins are thought to play a role in either alkane uptake, tolerance or directly in alkane degradation of this bacterium (Kato *et al*., 2009; Kato *et al*., 2010). The presence of these six genes in *G. stearothermophilus* Tok5a2 suggest the strain possesses the capacity for long-chain alkane degradation. Long-chain alkane degradation has previously been reported for *G. stearothermophilus* strains.

Genes associated with the degradation of the aromatic compound phenol. The genomes of all three of the studied strains incorporate a gene coding for a phenol 2-monooxygenase (PheA EC: 1.14.13.7), which is responsible for the conversion of phenol to catechol (Nurk *et al*., 1991). The orthologous proteins (DSM465^{T}_CDS2936, Ham3A8_CDS1001 and Tok5A2_CDS998) in the three *Geobacillus* strains share 100.00% (Bitscore 570; E-value: 3.00E-164), 99.65% (Bitscore: 566; E-value: 3.00E-163) and 91.81% (Bitscore 533; E-value: 3.00E-153) amino acid identity with PheA of *G. thermodenitrificans* NG80-2 (WP_011887902). The genome of *G. thermodenitrificans* Ham3A8 furthermore carries four co-localized genes, *pheCDEF* (Ham3A8_CDS314, Ham3A8_CDS311, Ham3A8_CDS313 and Ham3A8_CDS312), which are absent from the other two strains. The encoded proteins share 71.45% average amino acid identity with the plasmid bound PheCDEF proteins in *P. thermoglucosidasius* DSM 6285 (ACA01540, AAZ76889, ACA01539 and ACA01538) (Omokoko et al., 2008). These proteins represent a 4-oxalocrotonate decarboxylase (PheC), 2-oxopent-4-dienoate hydratase (PheD), 4-hydroxy-2-oxovalerate aldolase (PheE) and acetaldehyde dehydrogenase (PheF) respectively. These enzymes are responsible for the successive degradation of 4-oxalocrotonate to acetyl-CoA (Omokoko *et al*.,

2008). The presence of these genes on the genome of *G. thermodenitrificans* Ham3A8 suggest that the strain is capable of phenol utilization.

Table 3.5: Local Blast analysis results for proteins with known functions associated with long-chain alkane and aromatic degradation and biosurfactant production.

Protein	**Function**	**Organism**	**GenBank Accession #**	**DSM 465^T**	**Ham3A8**	**Tok5A2**
Alkane degradation associated proteins						
AlkA	Alkane 1 monooxygenase	*Nocardia cyriacigeorgica* GUH-2	CCF65099	-	-	-
AlkB	Alkane 1 monooxygenase	*Pseudomonas putida* P1	Q9WWW6	-	-	-
AlkB	Alkane 1 monooxygenase	*Geobacillus thermoleovorans* T70	Q6EV94	-	-	-
AlkB-geo1	Alkane 1 monooxygenase	*Geobacillus icigianus* G1w1	MG720135	-	-	-
AlkB-geo2	Alkane 1 monooxygenase	*Geobacillus subterraneus* DSM 13552^T	ABU48583	-	-	-
AlkB-geo3	Alkane 1 monooxygenase	*Geobacillus jurassicus* DSM 15726^T	ABU48586	-	-	-
AlkB-geo4	Alkane 1 monooxygenase	*Geobacillus icigianus* G1w1	AYK02781	-	-	-
AlkB-geo5	Alkane 1 monooxygenase	*Geobacillus jurassicus* DSM 15726^T	ABU48588	-	-	-

Protein	Function	Organism	GenBank Accession #	DSM 465^T	Ham3A8	Tok5A2
AlkB-geo6	Alkane 1 monooxygenase	*Geobacillus icigianus* G1w1	AYK02782	-	-	-
AlkB-geo7	Alkane 1 monooxygenase	*Geobacillus uzenensis* DSM 13551^T	ABU48577	-	-	-
AlkB-geo8	Alkane 1 monooxygenase	*Geobacillus stearothermophilus* DSM 22^T	ABU48546	-	-	-
AlkB1	Alkane monooxygenase-1	*Rhodococcus erythropolis* PR4	AJA79003	-	-	-
AlkB2	Alkane monooxygenase-2	*Rhodococcus erythropolis* PR4	AJA79004	-	-	-
AlkF	Rubredoxin-2	*Ps*eudomonas *putida* P1	Q9WWW5	-	-	-
AlkG	Rubredoxin-1	*Ps*eudomonas *putida* P1	Q9WWW4	-	-	-
AlkH	Aldehyde dehydrogenase	*Ps*eudomonas *putida* P1	Q9WWW3	-	-	-
AlkJ	Alcohol dehydrogenase	*Ps*eudomonas *putida* P1	Q9WWW2	-	-	-
AlkK	Acyl-CoA synthetase	*Ps*eudomonas *putida* P1	Q9L4M6	-	-	-
AlkL	Outer membrane protein	*Ps*eudomonas *putida* P1	Q9L4M5	-	-	-

Protein	Function	Organism	GenBank Accession #	DSM 465^{T}	Ham3A8	Tok5A2
AlkM	Alkane 1-monooxygenase	*Acinetobacter calcoaceticus* EB104	Q9XDP7	-	-	-
AlkMa	Alkane hydroxylase A	*Acinetobacter* sp. M-1	Q9AQK2	-	-	-
AlkMb	Alkane hydroxylase B	*Acinetobacter sp. M-1*	Q9AQJ9	-	-	-
AlkN	Methyl-accepting chemotaxis protein	*Ps*eudomonas *putida P1*	Q9R9U8	-	-	-
AlkS	Transcriptional regulator	*Ps*eudomonas *putida* P1	Q9L4M7	-	-	-
AlkT	Rubredoxin-NAD(+) reductase	*Ps*eudomonas *putida* P1	Q9L4M8	-	-	-
AlmA	FAD-binding monooxygenase	*Acinetobacter* sp. DSM 17874	A5H9N6	-	-	-
AhpG	*Cytochrome P450 CYP153 alkane hydroxylase*	*Mycobacterium sp. HXN-1500*	CAH04396	-	-	-
AhpG1	*Cytochrome P450 CYP153 alkane monooxygenase*	*Sphingomonas macrogoltabida HXN-200*	CAH61448	-	-	-

Protein	Function	Organism	GenBank Accession #	DSM 465^T	Ham3A8	Tok5A2
AhpG2	*Cytochrome P450 CYP153 alkane monooxygenase*	*Alcanivorax borkumensis AP1*	CAH59968	-	-	-
AhpG3	*Cytochrome P450 CYP153 alkane monooxygenase*	*Sphingomonas macrogoltabida HXN-200*	CAH61454	-	-	-
AhpG4	*Cytochrome P450 CYP153 alkane monooxygenase*	*Sphingomonas macrogoltabida HXN-200*	CAH61456	-	-	-
Bt-Aldh	Aldehyde dehydrogenase	*Geobacillus thermoleovorans* B23	Q9FAB1	-	-	+
LadA	*Alkane monooxygenase*	*Geobacillus thermodenitrificans NG80-2*	ABO68832	-	-	-
LadAα_{B23}	*FMN-dependent alkane monooxygenase*	*Geobacillus thermoleovorans B23*	BAM76377	-	-	+
LadAβ_{B23}	*FMN-dependent alkane monooxygenase*	*Geobacillus thermoleovorans B23*	BAM76372	-	-	+

Protein	**Function**	**Organism**	**GenBank Accession #**	**DSM 465^{T}**	**Ham3A8**	**Tok5A2**
$LadB_{B23}$	*FMN-dependent alkane monooxygenase*	*Geobacillus thermoleovorans B23*	BAM76371	-	-	+
NonM	*Cytochrome P450 alkane monooxygenase*	*Acinetobacter sp. EB104*	CAC37904	-	-	-
P16	*Membrane protein*	*Geobacillus thermoleovorans B23*	BAB16679	-	-	+
P21	*Membrane protein*	*Geobacillus thermoleovorans B23*	BAB16602	-	-	+
Aromatic hydrocarbon degradation associated proteins						
CatB	Muconate cycloisomerase 1	*Ps*eudomonas *putida* PRS2015	P08310	-	-	-
NdoB	Naphthalene 1,2-dioxygenase	*Ps*eudomonas *putida* DSM 8368	P0A110	-	-	-
PheA	*Prephenate dehydratase*	*Geobacillus thermodenitrificans NG80-2*	WP_011887902	+	+	+

Protein	Function	Organism	GenBank Accession #	DSM 465^{T}	Ham3A8	Tok5A2
PheA1	*Phenol-hydroxylase large subunit*	*Parageobacillus thermoglucosidasius* DSM 6285	Q45413	-	-	-
PheA2	Phenol-hydroxylase small subunit	*Parageobacillus thermoglucosidasius* DSM 6285	Q3YB14	-	-	-
PheB	Catechol 2,3-dioxygenase	*Parageobacillus thermoglucosidasius* DSM 6285	AAZ76884	-	-	-
PheC	4-oxalocrotonate decarboxylase	*Parageobacillus thermoglucosidasius* DSM 6285	ACA01540	-	+	-
PheD	2-hydroxypenta-2,4-dienoate-hydratase	*Parageobacillus thermoglucosidasius* DSM 6285	AAZ76889	-	+	-
PheE	4-hydroxy-2-oxovalerate aldolase	*Parageobacillus thermoglucosidasius* DSM 6285	ACA01539	-	+	-
PheF	Acetaldehyde dehydrogenase	*Parageobacillus thermoglucosidasius* DSM 6285	ACA01538	-	+	-

Protein	Function	Organism	GenBank Accession #	DSM 465^T	Ham3A8	Tok5A2
PheR	Transcriptional regulator	*Parageobacillus thermoglucosidasius* DSM 6285	AAZ76888	-	-	-
Biosurfactant production related proteins						
FenA	Fengycin synthetase	*Bacillus subtilis* F29-3	O30980	-	-	-
FenB	Fengycin synthetase	*Bacillus subtilis* F29-3	Q45563	-	-	-
FenC	Peptide synthetase	*Bacillus subtilis* F29-3	O87606	-	-	-
FenD	Fengycin synthetase	*Bacillus subtilis* F29-3	O87704	-	-	-
FenE	Fengycin synthetase	*Bacillus subtilis* F29-3	O30981	-	-	-
ItuA	Iturin A synthetase A	*Bacillus subtilis* RB14	Q93I56	-	-	-
ItuB	Iturin A synthetase B	*Bacillus subtilis* RB14	Q93I55	-	-	-
ItuC	Iturin A synthetase C	*Bacillus subtilis* RB14	Q93I54	-	-	-
ItuD	Putative malonyl-CoA transacylase	*Bacillus subtilis* RB14	Q93I57	-	-	-
LchAA	Lichenysin synthetase subunit A	*Bacillus licheniformis* BNP29	CAA06323	-	-	-
LchAB	Lichenysin synthetase subunit B	*Bacillus licheniformis* BNP29	CAA06324	-	-	-

Protein	Function	Organism	GenBank Accession #	DSM 465^{T}	Ham3A8	Tok5A2
LchAC	Lichenysin synthetase subunit C	*Bacillus licheniformis* BNP29	CAA06325	-	-	-
MycA	Mycosubtilin synthase subunit A	*Bacillus subtilis* subsp. *spizizenii* ATCC 6633	Q9R9J1	-	-	-
MycB	Mycosubtilin synthase subunit B	*Bacillus subtilis* subsp. *spizizenii* ATCC 6633	Q9R9J0	-	-	-
MycC	Mycosubtilin synthase subunit C	*Bacillus subtilis* subsp. *spizizenii* ATCC 6633	Q9R9I9	-	-	-
RhlA	Rhamnosyltransferase chain A	*P*seudomonas *aeruginosa* PSE305	B7VGG0	-	-	-
RhlB	Rhamnosyl transferase	*P*seudomonas *aeruginosa* DSM 2659	Q51560	-	-	-
RhlC	Rhamnosyltransferase 2	*Pseudomonas aeruginosa* ATCC 15692	Q9I4K5	-	-	-
SboA	Subtilosin-A	*Bacillus subtilis* subsp. *subtilis* str. 168	O07623	-	-	-
Sfp	Surfactin synthase-activating enzyme	*Bacillus subtilis* subsp. *subtilis* str. 168	P39135	-	-	-
SrfAA	Surfactin synthase subunit 1	*Bacillus subtilis* subsp. *subtilis* str. 168	P27206	-	-	-

Protein	Function	Organism	GenBank Accession #	DSM 465^{T}	Ham3A8	Tok5A2
SrfAB	Surfactin synthase subunit 2	*Bacillus subtilis* subsp. *subtilis* str. 168	Q04747	-	-	-
SrfAC	Surfactin synthase subunit 3	*Bacillus subtilis* subsp. *subtilis* str. 168	Q08787	-	-	-
SrfAD	Surfactin synthase thioesterase subunit	*Bacillus subtilis* subsp. *subtilis* str. 168	Q08788	-	-	-

3.4 Conclusions

In Chapter 2, a collection of 134 thermophiles was screened for their ability to degrade hydrocarbons. One strain *G. stearothermophilus* Tok5A2 showed extensive oil degradation in liquid assays. By contrast, two other *Geobacillus* strains, *G. thermodenitrificans* Ham3a8 and DSM 465^T, displayed poor hydrocarbon degradation capacities but showed interesting oil displacement activities in plate assays. Here we have undertaken 2D GC/MS to further elucidate differences in their hydrocarbon degradation profiles. Unexpectedly, degradation of C_{12-14} n-alkanes was observed in all three *Geobacillus* strains with *G. thermodenitrificans* DSM 465^T demonstrating C_{12-16} n-alkane degradation. 2D GC/MS analysis also identified substantial increases in C_{27} alkanes for *G. thermodenitrificans* DSM 465^T and Ham3A8 and C_{44} alkanes for *G. thermodenitrificans* DSM 465^T compared to the control. This was likely due to alkane biosynthesis by the evaluated bacterial strains. C_{27} biosynthesis in particular, has previously been reported for bacteria (Ladygina *et al*., 2006). Further chemical analyses of the produced compounds should be undertaken to elucidate the potential function of the synthesised alkanes.

We further sequenced the genomes of *G. thermodenitrificans* Ham3A8 and *G. stearothermophilus* Tok5A2 to complement the complete genome of *G. thermodenitrificans* DSM 465^T. RAST genome annotation identified genes involved in the degradation of several aromatic compounds, gentisate, salicylate and quinate, as well as for alkane sulfonate in all three strains. Furthermore, genes involved in the degradation of toluenesulfonate in *G. thermodenitrificans* DSM 465^T and Ham3A8, along with benzoate degradation in *G. thermodenitrificans* Ham3A8. Subsequent Blast analysis also identified PheACDEF orthologues encoded on the genome of *G. thermodenitrificans* Ham3A8, suggesting the capacity for phenol utilization.

The three genomes were screened for the presence of proteins with characterized roles in alkane degradation. While orthologues of AlkB have been observed in numerous bacterial genera, and are frequently identified in *Geobacillus* spp., none were identified in any of the three strains analysed. Genome Blast analysis of the

three strains identified alkane degrading associated genes only on the genome of *G. stearothermophilus* Tok5A2. Degradation of crude oil components has previously been observed by *G. stearothermophilus* strains, however, has been attributed to the presence of AlkB orthologues. Three alkane monooxygenases, $LadA\alpha_{B23}$, $LadA\beta_{B23}$ $LadB_{B23}$, and three alkane degradation associated orthologues of Bt-Aldh, P16 and P21, similar to those of *G. thermoleovorans* B23 are encoded on the genome of *G. stearothermophilus* Tok5A2. It is likely that the presence of these proteins encoded on the genome of *G. stearothermophilus* Tok5A2 were the cause of the strong oil degradation activity observed for this strain in Chapter 2 (2.4.1). No known alkane degradation associated proteins were found to be encoded on the genomes of *G. thermodenitrificans* DSM 465^T and Ham3A8. As such the observed degradation of $C_{12\text{-}14}$ alkanes by these stains was not attributable to any known enzymes and may indicate the presence of yet uncharacterised alkane degradation pathways in these strains. Further investigation is required to elucidate the mechanisms behind the observed degradation by these strains. Identification of upregulated genes in cultures of these bacteria grown on $C_{12\text{-}14}$ enriched minimal media may facilitate this process.

A significant number of strains screened in Chapter 2 (2.3.2) displayed oil displacement activities. The presence of a large C_{27} peak in the 2D GC/MS analysis may have potentially been the result of biosurfactant production associated with observed oil displacement activities. No biosurfactant producing genes associated with C_{27} production were identified in any of the strains, although predicted subtilin-like genes were identified on the genomes of all three strains. Further characterisation of potential biosurfactant production capacities of strains which displayed oil displacement activity may further elucidate the activities and C_{27} peaks observed.

G. stearothermophilus Tok5A2 appears to be a good candidate for further investigation as a petroleum hydrocarbon bioremediation agent based on its performance throughout all analyses conducted. While *G. thermodenitrificans* DSM 465^T and Ham3A8 did not possess any described long-chain alkane degrading genes, they demonstrated the capacity for $C_{12\text{-}16}$ and $C_{12\text{-}14}$ alkane degradation respectively by yet unidentified pathways. They further demonstrated the potential

for C_{27} n-alkane biosynthesis. While the strains may not make excellent candidates for alkane degradation applications, their capacities for aromatic hydrocarbon degradation and potential biosurfactant production may yet be further explored and characterised as they displayed interesting oil degradation and displacement activities.

3.5 References

Aliyu, H., Lebre, P., Blom, J., Cowan, D. and De Maayer, P., 2016. Phylogenomic re-assessment of the thermophilic genus *Geobacillus*. Systematic and Applied Microbiology, 39(8): 527-533.

Armacost, K., Musila, J., Gathiaka, S., Ellis, H.R. and Acevedo, O., 2014. Exploring the catalytic mechanism of alkanesulfonate monooxygenase using molecular dynamics. Biochemistry, 53(20): 3308-3317.

Auch, A.F., von Jan, M., Klenk, H.P. and Göker, M., 2010. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Standards in Genomic Sciences, 2(1): 117.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M. and Meyer, F., 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics, 9(1): 75.

Barbe, V., Cruveiller, S., Kunst, F., Lenoble, P., Meurice, G., Sekowska, A., Vallenet, D., Wang, T., Moszer, I., Médigue, C. and Danchin, A., 2009. From a consortium sequence to a unified sequence: the *Bacillus subtilis* 168 reference genome a decade later. Microbiology, 155: 1758.

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S.Y., Medema, M.H. and Weber, T., 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic acids research, 47(W1): W81-W87.

Boonmak, C., Takahashi, Y. and Morikawa, M., 2014. Cloning and expression of three *ladA*-type alkane monooxygenase genes from an extremely thermophilic alkane-degrading bacterium *Geobacillus thermoleovorans* B23. Extremophiles, 18(3): 515-523.

Bordoloi, N.K. and Konwar, B.K., 2009. Bacterial biosurfactant in enhancing solubility and metabolism of petroleum hydrocarbons. Journal of Hazardous Materials, 170(1): 495-505.

Bosi, E., Donati, B., Galardini, M., Brunetti, S., Sagot, M.F., Lió, P., Crescenzi, P., Fani, R. and Fondi, M., 2015. MeDuSa: a multi-draft based scaffolder. Bioinformatics, 31(15): 2443-2451.

Chen, S.H. and Aitken, M.D., 1999. Salicylate stimulates the degradation of high-molecular weight polycyclic aromatic hydrocarbons by *Pseudomonas saccharophila* P15. Environmental Science & Technology, 33(3): 435-439.

Cook, A.M., Laue, H. and Junker, F., 1998. Microbial desulfonation. FEMS Microbiology Reviews, 22(5): 399-419.

Dailey, H.A., Dailey, T.A., Gerdes, S., Jahn, D., Jahn, M., O'Brian, M.R. and Warren, M.J., 2017. Prokaryotic heme biosynthesis: Multiple pathways to a common essential product. Microbiology and Molecular Biology Reviews, 81(1): e00048-16.

Darling, A.C., Mau, B., Blattner, F.R. and Perna, N.T., 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Research, 14(7): 1394-1403.

Noudeh, G.D., Noodeh, A.D., Moshafi, M.H., Afzadi, M.A., Pardakhti, A. and Sal, M., 2010. Investigating the effects of various additives on surface activity and emulsification index of biosurfactant resulting from broth media of *Bacillus subtilis* PTCC 1023. African Journal of Microbiology Research, 4(19): 1981-1990.

Eichhorn, E., van der Ploeg, J.R. and Leisinger, T., 1999. Characterization of a two-component alkanesulfonate monooxygenase from *Escherichia coli*. Journal of Biological Chemistry, 274(38): 26639-26646.

Elumalai, P., Parthipan, P., Karthikeyan, O.P. and Rajasekar, A., 2017. Enzyme-mediated biodegradation of long-chain n-alkanes (C 32 and C 40) by thermophilic bacteria. 3 Biotech, 7(2), p.116.

Eppinger, E. and Stolz, A., 2016. Expansion of the substrate range of the gentisate 1, 2-dioxygenase from *Corynebacterium glutamicum* for the conversion of monohydroxylated benzoates. Protein Engineering, Design and Selection, 30(1): 57-65.

Feng, L., Wang, W., Cheng, J., Ren, Y., Zhao, G., Gao, C., Tang, Y., Liu, X., Han, W., Peng, X. and Liu, R., 2007. Genome and proteome of long-chain alkane degrading *Geobacillus thermodenitrificans* NG80-2 isolated from a deep-subsurface oil reservoir. Proceedings of the National Academy of Sciences, 104(13): 5602-5607.

Garg, N., Oman, T.J., Wang, T.S.A., De Gonzalo, C.V.G., Walker, S. and Van Der Donk, W.A., 2014. Mode of action and structure–activity relationship studies of geobacillin I. The Journal of Antibiotics, 67(1): 133.

Garg, N., Tang, W., Goto, Y., Nair, S.K. and van der Donk, W.A., 2012. Lantibiotics from *Geobacillus thermodenitrificans*. Proceedings of the National Academy of Sciences, 109(14): 5241-5246.

Giri, S.S., Sen, S.S., Jun, J.W., Sukumaran, V. and Park, S.C., 2017. Role of *Bacillus licheniformis* VS16-derived biosurfactant in mediating immune responses in Carp Rohu and its application to the food industry. Frontiers in Microbiology, 8: 514.

Grant, J.R. and Stothard, P., 2008. The CGView Server: a comparative genomics tool for circular genomes. Nucleic acids research, 36(2): W181-W184.

Grund, E., Denecke, B. and Eichenlaub, R., 1992. Naphthalene degradation via salicylate and gentisate by *Rhodococcus* sp. strain B4. Applied Environmental Microbiology, 58(6): 1874-1877.

Grund, E. and Kutzner, H.J., 1998. Utilization of quinate and p-hydroxybenzoate by actinomycetes: key enzymes and taxonomic relevance. Journal of Basic Microbiology: An International Journal on Biochemistry, Physiology, Genetics, Morphology, and Ecology of Microorganisms, 38(4): 241-255.

Guizhou, G., Zheng, L., Dongfeng, Z. and Chaocheng, Z., 2013. Isolation and characterization of a thermophilic oil degrading bacterial consortium. China Petroleum Processing and Petrochemical Technology, 15(2): 82-90.

Hall, T., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis. Nucleic Acids Symposium Series, 41: 95-98.

Hassanshahian, M., Emtiazi, G. and Cappello, S., 2012. Isolation and characterization of crude-oil-degrading bacteria from the Persian Gulf and the Caspian Sea. Marine Pollution Bulletin, 64(1): 7-12.

Hsiao, T.L., Revelles, O., Chen, L., Sauer, U. and Vitkup, D., 2010. Automatic policing of biochemical annotations using genomic correlations. Nature Chemical Biology, 6(1): 34.

Hussein, A.H., Lisowska, B.K. and Leak, D.J., 2015. The genus *Geobacillus* and their biotechnological potential. Advances in Applied Microbiology, 92: 1-48.

Jaggi, A., Snowdon, R.W., Stopford, A., Radović, J.R., Oldenburg, T.B. and Larter, S.R., 2017. Experimental simulation of crude oil-water partitioning behaviour of BTEX compounds during a deep submarine oil spill. Organic Geochemistry, 108: 1-8.

Jimenez_Diaz, L., Caballero, A. and Segura, A., 2019. Pathways for the degradation of fatty acids in bacteria. Aerobic Utilization of Hydrocarbons, Oils and Lipids, Handbook of Hydrocarbon and Lipid Microbiology, Rojo, F., Ed 1, 1-23.

Kahnert, A., Vermeij, P., Wietek, C., James, P., Leisinger, T. and Kertesz, M.A., 2000. The *ssu* Locus Plays a Key Role in Organosulfur Metabolism in *Pseudomonas putida* S-313. Journal of Bacteriology, 182(10): 2869-2878.

Kannenberg, E.L. and Poralla, K., 1999. Hopanoid biosynthesis and function in bacteria. Naturwissenschaften, 86(4): 168-176.

Kato, T., Miyanaga, A., Kanaya, S. and Morikawa, M., 2009. Alkane inducible proteins in *Geobacillus thermoleovorans* B23. BMC Microbiology, 9(1): 60.

Kato, T., Miyanaga, A., Kanaya, S. and Morikawa, M., 2010. Gene cloning and characterization of an aldehyde dehydrogenase from long-chain alkane-degrading *Geobacillus thermoleovorans* B23. Extremophiles, 14(1): 33.

Kiesau, P., Eikmanns, U., Gutowski-Eckel, Z., Weber, S., Hammelmann, M. and Entian, K.D., 1997. Evidence for a multimeric subtilin synthetase complex. Journal of Bacteriology, 179(5): 1475-1481.

Kleerebezem, M., Bongers, R., Rutten, G., De Vos, W.M. and Kuipers, O.P., 2004. Autoregulation of subtilin biosynthesis in *Bacillus subtilis*: the role of the spa-box in subtilin-responsive promoters. Peptides, 25(9): 1415-1424.

Klein, C. and Entian, K.D., 1994. Genes involved in self-protection against the lantibiotic subtilin produced by *Bacillus subtilis* ATCC 6633. Applied Environmental Microbiology, 60(8): 2793-2801.

Konz, D., Doekel, S. and Marahiel, M.A., 1999. Molecular and biochemical characterization of the protein template controlling biosynthesis of the lipopeptide lichenysin. Journal of Bacteriology, 181(1): 133-140.

Koumoutsi, A., Chen, X.H., Henne, A., Liesegang, H., Hitzeroth, G., Franke, P., Vater, J. and Borriss, R., 2004. Structural and functional characterization of gene clusters directing nonribosomal synthesis of bioactive cyclic lipopeptides in *Bacillus amyloliquefaciens* strain FZB42. Journal of Bacteriology, 186(4): 1084-1096.

Kuboi, R., Maruki, T., Tanaka, H. and Komasawa, I., 1994. Fermentation of *Bacillus subtilis* ATCC 6633 and production of subtilin in polyethylene glycol/phosphate aqueous two-phase systems. Journal of Fermentation and Bioengineering, 78(6): 431-436.

Ladygina, N., Dedyukhina, E.G. and Vainshtein, M.B., 2006. A review on microbial synthesis of hydrocarbons. Process Biochemistry, 41(5): 1001-1014.

Lee, I., Kim, Y.O., Park, S.C. and Chun, J., 2016. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. International Journal of Systematic and Evolutionary Microbiology, 66(2): 1100-1103.

Lefort, V., Desper, R. and Gascuel, O., 2015. FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. Molecular Biology and Evolution, 32(10): 2798-2800.

Marseille, F., Disnar, J.R., Guillet, B. and Noack, Y., 1999. n-Alkanes and free fatty acids in humus and A1 horizons of soils under beech, spruce and grass in the

Massif-Central (Mont-Lozère), France. European Journal of Soil Science, 50(3): 433-441.

Martínez, B., Böttiger, T., Schneider, T., Rodríguez, A., Sahl, H.G. and Wiedemann, I., 2008. Specific interaction of the unmodified bacteriocin Lactococcin 972 with the cell wall precursor lipid II. Applied Environmental Microbiology, 74(15): 4666-4670.

Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.P. and Göker, M., 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics, 14(1): 60.

Meier-Kolthoff, J.P. and Göker, M., 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nature Communications, 10(1): 2182.

Miethke, M., Westers, H., Blom, E.J., Kuipers, O.P. and Marahiel, M.A., 2006. Iron starvation triggers the stringent response and induces amino acid biosynthesis for bacillibactin production in *Bacillus subtilis*. Journal of Bacteriology, 188(24): 8655-8657.

Nešvera, J., Rucká, L. and Pátek, M., 2015. Catabolism of phenol and its derivatives in bacteria: genes, their regulation, and use in the biodegradation of toxic pollutants. Advances in Applied Microbiology, 93: 107-160.

Nurk, A., Kasak, L. and Kivisaar, M., 1991. Sequence of the gene (*pheA*) encoding phenol monooxygenase from *Pseudomonas* sp. EST1001: expression in *Escherichia coli* and *Pseudomonas putida*. Gene, 102(1): 13-18.

Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., Prjibelsky, A., Pyshkin, A., Sirotkin, A., Sirotkin, Y. and Stepanauskas, R., 2013, April. Assembling genomes and mini-metagenomes from highly chimeric reads. Annual International Conference on Research in Computational Molecular Biology, 7821: 158-170.

Omokoko, B., Jäntges, U.K., Zimmermann, M., Reiss, M. and Hartmeier, W., 2008. Isolation of the *phe*-operon from *G. stearothermophilus* comprising the phenol

degradative meta-pathway genes and a novel transcriptional regulator. BMC Microbiology, 8(1): 197.

Overbeek, R., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H.Y., Cohoon, M., de Crécy-Lagard, V., Diaz, N., Disz, T., Edwards, R. and Fonstein, M., 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic acids research, 33(17): 5691-5702.

Pao, S.S., Paulsen, I.T. and Saier, M.H., 1998. Major facilitator superfamily. Microbiology and Molecular Biology Reviews, 62(1): 1-34.

Peng, C., Huang, D., Shi, Y., Zhang, B., Sun, L., Li, M., Deng, X. and Wang, W., 2019. Comparative transcriptomic analysis revealed the key pathways responsible for organic sulfur removal by thermophilic bacterium *Geobacillus thermoglucosidasius* W-2. Science of the Total Environment, 676: 639-650.

Perez, K.J., Viana, J.D.S., Lopes, F.C., Pereira, J.Q., dos Santos, D.M., Oliveira, J.S., Velho, R.V., Crispim, S.M., Nicoli, J.R., Brandelli, A. and Nardi, R., 2017. *Bacillus* spp. isolated from puba as a source of biosurfactants and antimicrobial lipopeptides. Frontiers in Microbiology, 8: 61.

Phale, P.S., Basu, A., Majhi, P.D., Deveryshetty, J., Vamsee-Krishna, C. and Shrivastava, R., 2007. Metabolic diversity in bacterial degradation of aromatic compounds. Omics: A Journal of Integrative Biology, 11(3): 252-279.

Pradhan, A.K., Pradhan, N., Sukla, L.B., Panda, P.K. and Mishra, B.K., 2014. Inhibition of pathogenic bacterial biofilm by biosurfactant produced by *Lysinibacillus fusiformis* S9. Bioprocess and Biosystems Engineering, 37(2): 139-149.

Sáenz, J.P., Grosser, D., Bradley, A.S., Lagny, T.J., Lavrynenko, O., Broda, M. and Simons, K., 2015. Hopanoids as functional analogues of cholesterol in bacterial membranes. Proceedings of the National Academy of Sciences, 112(38): 11971-11976.

Seaver, L.C. and Imlay, J.A., 2001. Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in Escherichia coli. Journal of bacteriology, 183(24), pp.7173-7181.

Shestakova, N.M., Korshunova, A.V., Mikhailova, E.M., Sokolova, D.S., Tourova, T.P., Belyaev, S.S., Poltaraus, A.B. and Nazina, T.N., 2011. Characterization of the aerobic hydrocarbon-oxidizing enrichments from a high-temperature petroleum reservoir by comparative analysis of DNA-and RNA-derived clone libraries. Microbiology, 80(1): 60-69.

Speight, J.G., 2014. The chemistry and technology of petroleum. 5th ed. Boca Raton, FL. CRC press.

Stein, T., Heinzmann, S., Kiesau, P., Himmel, B. and Entian, K.D., 2003. The spa-box for transcriptional activation of subtilin biosynthesis and immunity in *Bacillus subtilis*. Molecular Microbiology, 47(6): 1627-1636.

Tourova, T.P., Nazina, T.N., Mikhailova, E.M., Rodionova, T.A., Ekimov, A.N., Mashukova, A.V. and Poltaraus, A.B., 2008. *alkB* homologs in thermophilic bacteria of the genus *Geobacillus*. Molecular Biology, 42(2): 217.

Valdes-Stauber, N. and Scherer, S., 1996. Nucleotide sequence and taxonomical distribution of the bacteriocin gene *lin* cloned from *Brevibacterium linens* M18. Applied Environmental Microbiology, 62(4): 1283-1286.

van Beilen, J.B., Wubbolts, M.G. and Witholt, B., 1994. Genetics of alkane oxidation by *Pseudomonas oleovorans*. Biodegradation, 5(4): 161-174.

van der Ploeg, J.R., Iwanicka-Nowicka, R., Bykowski, T., Hryniewicz, M.M. and Leisinger, T., 1999. The *Escherichia coli ssuEADCB* gene cluster is required for the utilization of sulfur from aliphatic sulfonates and is regulated by the transcriptional activator *Cbl*. Journal of Biological Chemistry, 274(41): 29358-29365.

Vanittanakom, N., Loeffler, W., Koch, U. and Jung, G., 1986. Fengycin-a novel antifungal lipopeptide antibiotic produced by *Bacillus subtilis* F-29-3. The Journal of Antibiotics, 39(7): 888-901.

Varjani, S.J. and Upasani, V.N., 2017. A new look on factors affecting microbial degradation of petroleum hydrocarbon pollutants. International Biodeterioration & Biodegradation, 120: 71-83.

Wang, L., Tang, Y., Wang, S., Liu, R.L., Liu, M.Z., Zhang, Y., Liang, F.L. and Feng, L., 2006. Isolation and characterization of a novel thermophilic *Bacillus* strain degrading long-chain n-alkanes. Extremophiles, 10(4): 347.

Wentzel, A., Ellingsen, T.E., Kotlar, H.K., Zotchev, S.B. and Throne-Holst, M., 2007. Bacterial metabolism of long-chain n-alkanes. Applied Microbiology and Biotechnology, 76(6): 1209-1221.

Zhang, Z., Schwartz, S., Wagner, L. and Miller, W., 2000. A greedy algorithm for aligning DNA sequences. Journal of Computational biology, 7(1-2): 203-214.

Zheng, B., Zhang, F., Dong, H., Chai, L., Shu, F., Yi, S., Wang, Z., Cui, Q., Dong, H., Zhang, Z. and Hou, D., 2016. Draft genome sequence of *Paenibacillus* sp. strain A2. Standards in Genomic Sciences, 11(1): 9.

Zhou, J.F., Gao, P.K., Dai, X.H., Cui, X.Y., Tian, H.M., Xie, J.J., Li, G.Q. and Ma, T., 2018. Heavy hydrocarbon degradation of crude oil by a novel thermophilic *Geobacillus stearothermophilus* strain A-2. International Biodeterioration & Biodegradation, 126: 224-230.

Ziemert, N., Alanjary, M. and Weber, T., 2016. The evolution of genome mining in microbes–a review. Natural Product Reports, 33(8): 988-1005.

SUMMARY

The global distribution and use of crude oil and petroleum derivatives facilitates the release of hundreds of millions of litres of these substances into vulnerable environments. With such vast volumes of these substances being released continually, remediation efforts to prevent long term environmental damage are imperative. Bioremediation has been identified as an environmentally innocuous and effective remediation technique. While application of mesophilic bacterial taxa in bioremediation approaches have been studied extensively, investigation into use of thermophilic bacteria has remained relatively untouched. Thermophilic bacteria, such as members of the genus *Geobacillus*, have been known to possess the capacity for the degradation of a range of diverse compounds, including those found in crude oil.

Here we have evaluated a collection of 134 thermophilic bacteria for the ability to degrade motor oil. A total of 58 strains were identified to show some capacity to degrade used-motor oil. These included taxa from the genera *Bacillus, Aeribacillus, Anoxybacillus* and *Geobacillus*. Furthermore, 63 strains demonstrated notable oil displacement. Culture age was noted to have variable effects on oil degradation capacities of the strains tested, with younger cultures having in general, higher oil degradation activities.

Three strains belonging to the genus *Geobacillus* were selected for further analysis. These included *G. thermodenitrificans* DSM 465^T and Ham3A8, which displayed poor/no apparent oil degradation capacities, but showed motor oil concentration activity. The third strain, *G. stearothermophilus* Tok5A2 showed no oil displacement activity but demonstrated strong oil degradation activity. 2D GC/MS analysis identified $C_{12\text{-}14}$ n-alkane degradation occurring in all three strains with $C_{12\text{-}16}$ n-alkane degradation by the *G. thermodenitrificans* DSM 465^T. Along with n-alkane degradation, substantial C_{27} (*G. thermodenitrificans* DSM 465^T and Ham3A8) and C_{44} (*G. thermodenitrificans* DSM 465^T) n-alkane biosynthesis was observed.

Comparative genome analysis of the three strains elucidated several pathways with potential roles in hydrocarbon degradation. Genes associated with the degradation

of alkanesulfonates and aromatic compounds, gentisate, salicylate and quinate were identified on the genomes of all three strains. Further genes involved in the degradation of toluenesulfonate (*G. thermodenitrificans* DSM 465^T and Ham3A8) benzoate and phenol (*G. thermodenitrificans* Ham3A8) were also identified.

No AlkB homologues were found to be encoded on the genomes of any of the evaluated strains. Alkane degradation by *G. stearothermophilus* Tok5A2 could, in part, be attributed to the presence of three alkane monooxygenases, LadAα_{B23}, LadAβ_{B23} and LadB$_{B23}$, and genes coding for orthologues of the alkane degradation associated proteins Bt-Aldh, P16 and P21 on the genome of the strain. No alkane degradation associated protein encoding genes were identified on the genomes of *G. thermodenitrificans* DSM 465^T and Ham3A8, even though GC/MS analysis identified $C_{12\text{-}16}$ and $C_{12\text{-}14}$ alkane degradation respectively. Further investigation will be required to characterise the pathways responsible for these degradation activities.

G. stearothermophilus Tok5A2 demonstrated consistent degradation activities in Chapter 2 and 3, along with possessing several alkane monooxygenases and alkane degradation associated proteins encoded on the genome of the strain. These results suggest that the strain is a good candidate for thermophilic n-alkane bioremediation applications. While *G. thermodenitrificans* DSM 465^T and Ham3A8 did show $C_{12\text{-}16}$ and $C_{12\text{-}14}$ alkane degradation respectively, their poor performance in the screening assays conducted in Chapter 2 and their respective production of C_{27} and C_{44} suggest they are not ideal candidates for thermophilic n-alkane bioremediation applications. However, both strains displayed interesting oil concentration activities which warrant further investigation.