Investigation of the total petroleum hydrocarbon degrading microorganisms in soil and water: a metagenomic approach

Firouz Abbasian

Submitted in total fulfilment of the requirements for the degree of Doctor of Philosophy

Global Centre for Environmental Remediation
Faculty of Science and Information Technology
The University of Newcastle

September 2016

DECLARATION

I declare that:

The thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. I give consent to the final version of my thesis being made available worldwide when deposited in the University's Digital Repository, subject to the provisions of the Copyright Act 1968.

Firouz Abbasian

Signed_____ Date: 9/9/2016

_____

ACKNOWLEDGEMENTS

First and foremost, I wish to thank my supervisor, Professor Megharaj Mallavarapu, who showed incredible patience and passion to supervise me throughout my study. I appreciate all attentions and support he devoted to me. At the same time, I thank my co-supervisor, Professor Ravi Naidu, who in addition to his spiritual support, have successfully managed the research team in both GCER and CRC-CARE, including me, to lead the researches in a productive and dynamic direction. Furthermore, I would like to appreciate Dr Robin Lockington, University of South Australia, for all his support and guides throughout my lab works.

Most importantly, I appreciate all my family, especially Azita, Adro and my parents, for all their support and encouragement throughout this period of study. I owe them for all their patience and spiritual support.

I appreciate all the staff and students in GCER and CRC-CARE, who have assisted me anytime I needed their help. Because of creation of a nice and warm environment they have made in these institutes, there was like home for me where I feel safety and happiness. Last but not least, I appreciate the University of Newcastle and the government of Australia for the admission offer, APA scholarship and IPRS scholarship, which enabled me to reach this position.

List of Publications:

1) F Abbasian, R Lockington, T. Palanisami, M Mallavarapu, R Naidu. (2016), Identification of a new operon involved in desulfurization of Dibenzothiophenes using a metagenomic study and cloning and functional analysis of the genes. Enzyme and Microbial Technology; 87: 24-28.
2) F Abbasian, R Lockington, T. Palanisami, K. Ramadass, M Mallavarapu, R Naidu (2016). Microbial diversity and hydrocarbon degrading gene capacity of a crude oil field soil as determined by metagenomics analysis. Biotechnology Progress. Accepted.
3) F Abbasian, R Lockington, T. Palanisami, M Mallavarapu, R Naidu. (2016), The biodiversity changes in the microbial population of soils contaminated with crude oil. Current Microbiology. 72: 663-670.
4) F Abbasian, R Lockington, T. Palanisami, M Mallavarapu, R Naidu. (2016). A review on the Genetics of Aliphatic and Aromatic Hydrocarbon Degradation. Applied biochemistry and biotechnology; 178 (2): 224-250.
5) F Abbasian, R Lockington, T. Palanisami, M Mallavarapu, R Naidu. (2015). Multiwall carbon nanotubes increase the microbial community in crude oil contaminated fresh water sediments. Accepted (Science of total environment).
6) F Abbasian, R Lockington, M Mallavarapu, R Naidu. (2015). The integration of sequencing and bioinformatics in metagenomics. Reviews in Environmental Science and Bio/Technology; 14(3): 357-383.
7) F Abbasian, R Lockington, M Mallavarapu, R Naidu. (2015). A pyrosequencing-based analysis of microbial diversity governed by ecological conditions in the Winogradsky column. World Journal of Microbiology and Biotechnology; 31(7): 1115-1126.
8) F Abbasian, R Lockington, M Mallavarapu, R Naidu. (2015). A Comprehensive Review of Aliphatic Hydrocarbon Biodegradation by Bacteria. Applied biochemistry and biotechnology; 176(3): 670-699.

TABLE OF CONTENTS

Abstract

Hydrocarbons are relatively recalcitrant compounds and are classified as high priority pollutants. However, these compounds are slowly degraded by a large variety of aerobic and anaerobic microorganisms. Although the corresponding genes in many phylogenetic groups of microbial species show different levels of diversity in terms of the gene sequence, the organisation of the genes in the genome or on plasmids and the activation mode of several microorganisms show identical hydrocarbon degrading genes.

Since the majority of microorganisms in natural environments cannot be cultured in laboratory media, culture-based systems are unable to estimate the full microbial diversity of an environment. Metagenomic methods, however, employ sequencing procedures for the determination of the microbial diversity of a community and for examining a particular functional ability of microorganisms in the environment using genomic DNA obtained directly from environmental samples. Application of metagenomic methods provides a huge amount of data that can be analysed only by using powerful computational bioinformatics tools.

In this study, we used next generation technology and metagenomic analysis to investigate the microbial diversity in crude oil and crude oil contaminated soils and to find the functional genes involved in the degradation of hydrocarbons in crude oil. The findings from this study can be used for bioremediation of crude oil spills and also for improvement of the quality of crude oil derivatives in terms of removal of sulfur and nitrogen. As a part of this study, we report a list of microorganisms that are abundant in the crude oil and the crude oil contaminated soil. Furthermore, we found a new operon responsible for removal of sulfur from dibenzothiophenes. The three genes in this operon were cloned and their activities measured in cell free condition.

Chapter 1 Introduction

Environmental contamination through the release of crude oil and its products is one of the common global environmental issues. The increasing demands for oil products in this century have led to an intensification of environmental pollution (Hao et al., 2004). Crude oil consists of a mix of various hydrocarbons, including alkanes and aromatics, which are categorized as serious and hazardous environmental pollutants (Nolvak et al., 2012, Anthony, 2006). The degree of toxicity of crude oil and its products depends on the chemical compounds present and their concentration. In addition, ecological niche and biological diversity affects toxicity. In the other words, different biological species, including both eukaryotes and prokaryotes, and their life stages respond differently to the pollutants (Hao et al., 2004).

Although all adverse aspects of oil pollutants on microbial diversity are not clearly known, it depends on the oil chemical composition and microbial species of the habitat (Sathishkumar et al., 2008). While some bacteria are killed directly by crude oil and/or its products, some others have a neutral interaction and some microbial populations are able to use oil components as their nutritional source. Indeed, these compounds are consumed as nutrients by such bacteria directly or after enzymatic digestion (Anthony, 2006). These microorganisms are useful for biological removal of crude oil and/or its products from contaminated sites, in a process named Bioremediation. Bioremediation is a way of cleaning up of polluted environments through the use of microbial metabolism to convert toxic contaminants into safe or less hazardous compounds or to convert oil compounds into biomass (Boopathy, 2000). As part of the oil polluted site-recovery process, microbial activities play fundamental roles in natural attenuation. Natural attenuation is a natural (biological or physical) procedure that reduces the toxicity and mass of the contaminants. It is clearly obvious that human activities, in terms of encouraging active bioremediation can facilitate the oil removal process (Anthony, 2006). Although oil-degrading microbes are ubiquitous, and come from many different microbial phyla (bacteria, archaea, fungi and algae), their ability to consume hydrocarbons is varied. While yeasts and bacteria, for instance, effectively degrade light chain hydrocarbons, they are less efficient in the case of higher chain compounds (Nolvak et al., 2012, Boopathy, 2000). Because microbial activity is necessary for oil bioremediation, characterization of microbial communities in the crude oil and in polluted environments is beneficial (Mail et al., 2012). This information helps us to identify the active members of an oil-utilising community and its probable changes in response to environmental parameters and specific oil substrates. This

information finally assists scientists to set up optimized chemical, physical and environmental conditions that may improve biodegradation of oil and its products. Although isolation of the majority of oil utilising microorganisms is not easy, application of molecular techniques enables researchers to improve their knowledge of microbial diversity in a variety of complex environments. Also, the molecular approaches may assist us to establish the correlations between the oil components and microbial diversity (Head et al., 2006).

In addition to bioremediation, the hydrocarbon degrading microorganisms may be used for improvement of the quality of crude oil and its derivatives in terms of removal of sulfur from both aromatic and aliphatic hydrocarbons (Masnadi et al., 2014). The combustion of these types of hydrocarbons leads to the production and release of sulfur oxides, which are considered as the main source for the formation of acid deposition (Khan et al., 2009). Removal of organic sulfur containing compounds is expensive, and design of biological system for removal of this element can be cost effective (Soleimani et al., 2007, Lu et al., 2007).

In this study we used a newly emerged technology, metagenomic study, firstly for looking for the presence of potential new microorganisms and the genes that are able to degrade hydrocarbons. Also, this technology enables us to find new catabolic genes, whose products are able to break down oil components. Furthermore, this study aimed to investigate the presence of the microorganisms and genes involved in improvement of crude oil in terms of removal of sulfur from hydrocarbons, and clone the interesting enzymes in an easily-handling bacterium, such as *E.coli*, for enzyme studies.

Chapter 2. Literature review

2.1 A Comprehensive Review of Aliphatic Hydrocarbon Biodegradation by Bacteria

Published in: Applied biochemistry and biotechnology; 176(3): 670-699.

2.1.1. Introduction

Hydrocarbons originate both from the biosynthetic activity of microorganisms and plants (as a result of enzymatic reduction of fatty acid molecules) (Schirmer et al., 2010) and from slow geochemical processes acting on biological compounds under high temperature and pressure over prolonged geological periods (McCarthy and Calvin, 2008). However increased demands for oil products due to industrialization have greatly increased the background level of exposure of the environment to different hydrocarbons (Odell, 2013). All the activities in the oil industries, including exploration, extraction, transportation, refining and oily waste management, are possible sources of environmental pollution (Hu et al., 2013). The effluents of petroleum wells, oil refinery operations, industrial waste waters, fuel consumption, wood-processing activities, detergents, pesticides, paints and other chemicals release huge quantities of hydrocarbons into soil and air (Rosenberg, 2013). Crude oils contain a mixture of different low and high molecular weight aliphatic hydrocarbons and several types of monocyclic and polycyclic aromatic compounds, in which one or more of the carbon atoms in their backbone can also be substituted by nitrogen, sulphur or oxygen. A soap molecule is also a hydrocarbon, consisting of a long hydrocarbon chain containing a carboxylic acid group at the end , and is the potassium or sodium salt of such a fatty acid (GUHA and GUPTA, 2013). The length and nature of hydrocarbons varies in oil derived compounds from C1-C10, such as make up Gasoline (Jimeson et al., 2012), C9 – C16, such as are found in Jet fuel and Kerosene (Chou et al., 2002) and C10-C50, such as waste oil (Monserud and Schwartz, 2012). The categories of typical hydrocarbon chains will be discussed later in the text. Since the components of hydrocarbons are highly recalcitrant to degradation (due to their low water solubility), they remain for a long time. The US EPA has classified these compounds as priority contaminants of natural resources (Husain, 2008). On the other hand, metabolism of hydrocarbons by microorganisms is a relatively complex process in which the organisms first take up the hydrocarbons and then convert these metabolically inactive molecules to more active forms for further catalysis (Ladygina et al., 2006). Although each group of hydrocarbons is broken down through specific enzymatic cascades, their final product is frequently similar.

This review attempts to describe the microbial abilities and the pathways used for degradation of hydrocarbons in both aerobic and anaerobic conditions and to update previous data using that from the latest publications. This review has focused more attention on the aliphatic hydrocarbons, although in some cases aromatic hydrocarbons were used as prototype examples.

2.1.2. Uptake of hydrocarbons into microbial cells

Microorganisms are challenged by the hydrophobicity and insolubility of hydrocarbons which both causes toxic effects on the cells, and restricts their absorption into the cells (Ben Ayed et al., 2013). Since hydrocarbons interrupt membrane fluidity, the cell membrane composition of the exposed microorganisms has to change in order to stabilize the structure of the phospholipid bilayer (Jin et al., 2014). The intensification of these changes depends on the type of hydrocarbons and their carbon chain length, and include changes from cis-to-trans isomers and from saturated fatty acids to unsaturated form or the reverse (Segura et al., 2012). For instance, while C2-C4 alcohols increase the ratio of unsaturated fatty acids in cell membrane, longer alkahols induce the production of saturated fatty acids (Segura et al., 2012, Jin et al., 2014). These limitations can be resolved using microdroplets, macrodroplets or dissolution of the hydrocarbon molecules into water (Bustamante et al., 2012). Most of the hydrocarbon degrading bacteria produce surfactant compounds in order to emulsify hydrocarbon molecules to form droplets or micelles, which finally are taken up by microorganisms in various ways (Bustamante et al., 2012). The extremely hydrophobic surfaces of many of the hydrocarbon degrading bacteria enable them to associate with hydrocarbon droplets or even to pass through into the organic phase of the droplets. Due to the difficulty of solubilisation and the slow dissolution of high molecular weight (HMW) hydrocarbons, their uptake in aquatic systems is slower than low molecular weight (LMW) hydrocarbons. Since the rate of metabolism of a hydrocarbon is dependent on its uptake, HMW hydrocarbons are metabolized slower than LMW hydrocarbons (Miller and Bartha, 1989). Both passive and active transportation systems for hydrocarbons have been found in microorganisms (Hua et al., 2013, Li et al., 2014). For instance, at a concentration of higher than 4.54 μmol/L *Pseudomonas* sp. DG17 takes up *n*-octadecane by a facilitated passive transport systems, while at lower concentrations of this hydrocarbon the system is switched into an energy-dependent transportation (Hua et al., 2013).

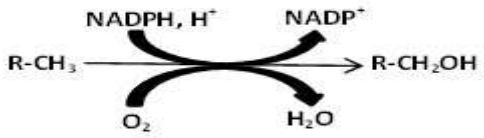
2.1.3. Metabolic shunts of n-Alkanes

n-Alkanes are the most swiftly decomposing components of aliphatic hydrocarbons, and are divided into four molecular weight groups: the gaseous alkanes, the aliphatic hydrocarbons with lower molecular weight (C8–C16), the aliphatic hydrocarbons with medium molecular weight (C17– C28) and finally, the aliphatic hydrocarbons with high molecular weight (>C28) (Riazi, 2005). Although the smallest group of hydrocarbons are easily degraded by most hydrocarbon degrading microorganisms, degradation of aliphatic hydrocarbons up to C44 also occurs. However, these long chain alkanes are first enzymatically activated before degradation (Table 2.1.1) (van Beilen and Funhoff, 2007).

Enzymes	Methane monooxygenases	Alkane monooxygenases	Bacterial P450 (CY153, Class I)	Eukaryotic P450 (CYP52, Class II)	Dioxygenases
Chain length	C1-C4	C5-C16	C5-C16	C10-C16	C10-C30

Table 2.1.1 The known enzymes involved in degradation of aliphatic hydrocarbons

First and most commonly, n-alkanes are oxidized by an electron carrier dependent monooxygenase system, the alkane hydroxylases, to their corresponding alkan-1-ol (Fig. 2.1.1) (Li et al., 2008b, Xie et al., 2011). Several bacteria, including toluene-oxidizing bacteria (Tinberg et al., 2011), ammonia-oxidizing bacteria (Taylor et al., 2013), methane oxidizing bacteria (Luesken et al., 2011) and propane-oxidizing bacteria (Redmond et al., 2010) can use monooxygenases to oxidize hydrocarbons by a reaction in which an oxygen atom from O_2 is incorporated into the hydrocarbons (Reaction 2.1). Based on the substrate, these enzymes are called toluene mono/dioxygenase, methane monooxygenase, ammonia monooxygenase, propane monooxygenase and alkane oxygenase (Tinberg et al., 2011, Taylor et al., 2013, Luesken et al., 2011, Redmond et al., 2010). In addition, many of these enzymes are broadly able to oxidise other types of hydrocarbons, such as halogenated aliphatics.



(Reaction 2.1)

The monooxygenases isolated in prokaryotes are classified into two categories based on their electron transport system and the microorganisms in which they are found: a) A rubredoxin dependent enzyme (containing 2FeO), encoded by the gene alkB in most of bacteria and alkM in *Acinetobacter* sp., b) an alkane hydroxylase containing cytochrome P450 monooxygenases in the CYP153 family of bacteria (Urlacher and Girhard, 2012, Morikawa, 2010, van Beilen and Funhoff, 2007). The first enzyme isolated was a non-heme diiron monooxygenase alkane hydroxylase located in the cell membrane of *Pseudomonas putida* GPo1 that is encoded by the gene alkB and requires the electron transporter subunit, rubredoxin reductase and rubredoxin (van Beilen et al., 1994). Later, other homologous enzymes with a conserved sequence at the active site of the protein were isolated from different microorganisms, especially α-Proteobacteria, β-Proteobacteria and γ-Proteobacteria as well as the Actinomycetales (Van Beilen et al., 2003). Except for some enzymes isolated from pseudomonads and some γ-Proteobacteria such as *A. borkumensis*, capable of mineralizing C5–C12 alkanes, other kinds of enzymes are required to attack aliphatic hydrocarbons longer than 10 atom carbons (van Beilen and Funhoff, 2007). LadA is a flavoprotein dependent monooxygenase isolated from a thermophilic microorganism (*Geobacillus thermodenitrificans* NG80-2) that activates the long-chain alkanes (C15 to C36) for degradation through breaking down an O_2 molecule and inserting an oxygen atom into its structure (Palfey et al., 1995). In this reaction, the flavin is reduced by NAD(P)H and then interact with O_2 to transfer one oxygen atom to the alkane (Li et al., 2008b).

The cytochrome P450 enzymes are a group of heme (iron protoporphyrin IX) containing monooxygenase enzymes found in both eukaryotes and prokaryotes that function in association with molecular oxygen, NAD(P)H and an electron-transfer system(s) to oxidize a large variety of compounds (Urlacher and Eiben, 2006). In contrast to eukaryotes, in which the enzyme is associated with cellular membranes, the bacterial P450 cytochromes are soluble in the cytoplasm (Urlacher and Girhard, 2012, Urlacher and Eiben, 2006). Based on the bacterial species, the structures of these enzymes are different and while in many of the bacteria these enzymes consist of three components (hydroxylase, P450cam and putidaredoxin), an analogous enzyme in *Bacillus megaterium (*P450BM3) contain only one large peptide-chain

that performs both the electron transportation and substrate oxygenation (Urlacher and Girhard, 2012). The active site of all categories of Cyt-P450 consists of a cysteinate and an iron protoporphyrin IX, and this interacts with O_2 for the oxidation of substrates (Gricman et al., 2013).

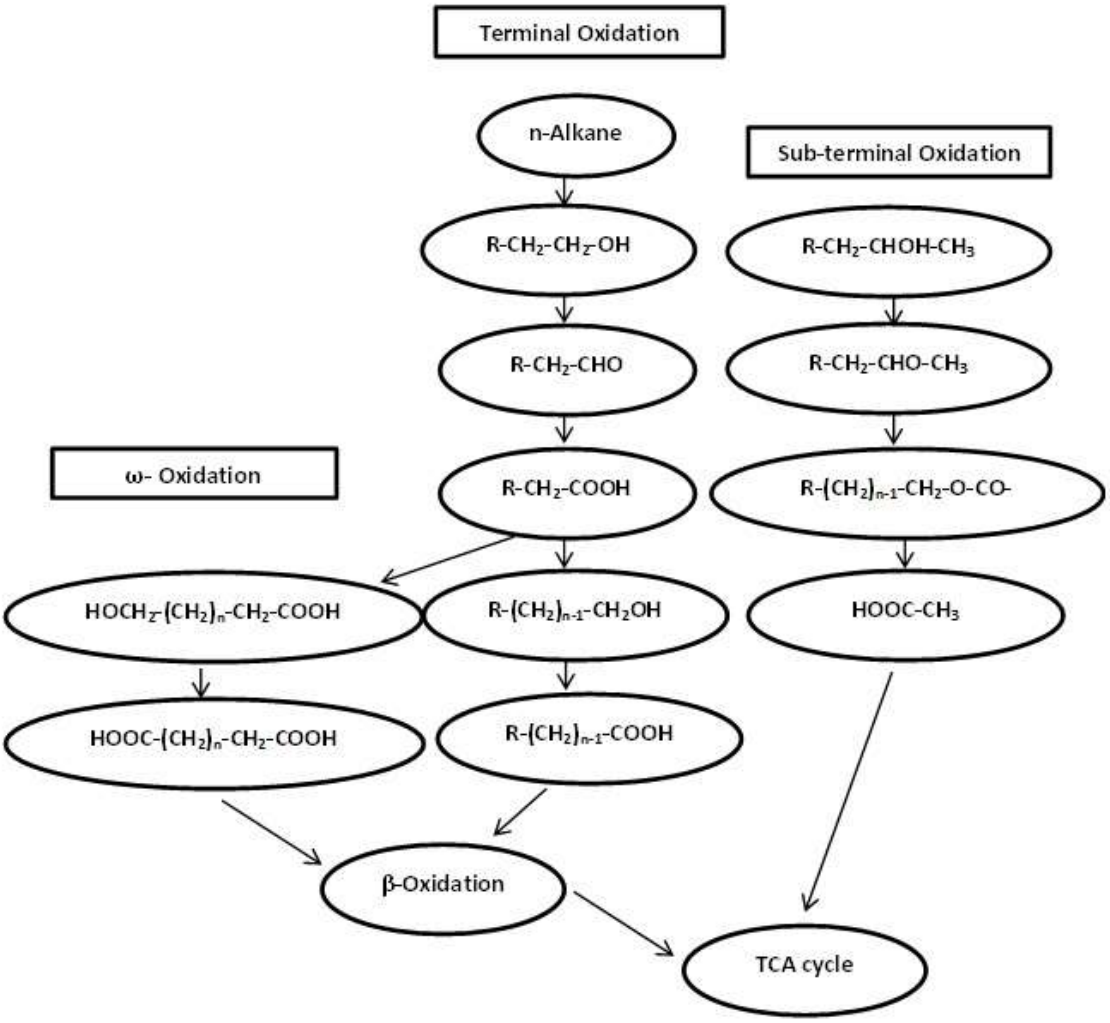


Figure 2.1.1 Three possible n-Alkane degradation pathways

Overall, the electron obtained from NADPH is transferred through an electron transport system, consisting of a reductase and FES redoxin compound, to the Cyt-P450 where the electrons are used for the hydroxylation of a substrate, such as a hydrocarbon, at the expense of breaking down molecular oxygen (Reaction 2.2) (Ortiz de Montellano, 2010). However, the cytochrome produced by *Bacillus megaterium* and *Sphingomonas paucimobilis* uses peroxides, instead of O_2 as the oxidant agent in the oxidative reaction (Ortiz de Montellano, 2010). Several

bacteria, such as *Rhodococcus rhodochrous* ($P450_{oct}$) (Larkin et al., 2005), *Acinetobacter* sp. (Throne-Holst et al., 2007) and many others utilize their Cyt P-450 for the initial activation of hydrocarbon degradation, yielding the corresponding alcohols (Kubota et al., 2005).

(Reaction 2.2) NADPH → Reductase (FAD and FMN) → Redoxin (FeS) → Cytochrome P450

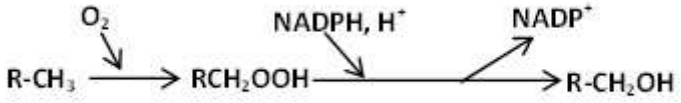
$$RH + O_2 \xrightarrow{\text{Cyt-P450}} ROH + H_2O$$

Overall: $RH + NAD(P)H + O_2 + H^+ \rightarrow ROH + NAD(P)^+ + H_2O$

In eukaryotes, including fungi, endoplasmic nucleotide-dependent cytochrome P450 reductases (CPR) transfer two electrons from NADPH to cytochrome P450 which uses these electrons for a variety of molecular biotransformation reactions such as the detoxification or biosynthesis of secondary compounds (Črešnar and Petrič, 2011). For example this reaction, in *Saccharomyces cerevisiae* is used in the ergosterol biosynthesis pathway and for the metabolism of many polyaromatic hydrocarbons such as benzo[a]pyrene (Črešnar and Petrič, 2011).

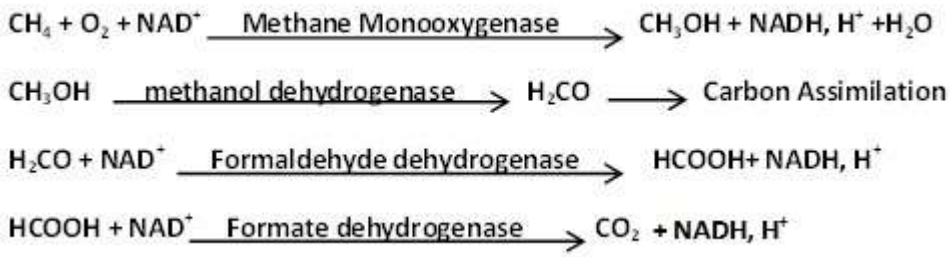
In addition to these two prototype reactions for the oxidation of alkanes, there are some other types of alkane metabolism which can be seen in special microorganisms. Through an alternative terminal pathway some yeasts such as *Mortierella* sp. utilize alcohol oxidases instead of alcohol dehydrogenases, producing either or both ω-hydroxic acids and α-ω,dioic acids (Fig. 2.1.1) (Anderson et al., 2008, Waché, 2013). Also, some bacteria such as *Bacillus* sp. and *Rhodococcus* sp. and some fungi such as *Fusarium* sp. and *Aspergillus* sp. are able to oxidize sub-terminally the n-alkanes to the 4-, 5- or 6-ol products and to a lesser amount the 2- and 3-ols. These alcohols are further oxidized by (pyridine nucleotide-linked) dehydrogenases into their corresponding aldehyde and fatty acids. The resulting fatty acids may enter a β-Oxidation process or may be inserted into the cell membrane (Binazadeh et al., 2009). In addition to these oxidative reactions in which only one atom oxygen is used, the Finnerty pathway is a process in which dioxygenase systems are able to transform n-alkanes first into their corresponding hydroperoxides and then into the corresponding alkan-1-ol (Reaction 2.3) (Ji et al., 2013). In this pathway, microorganisms oxidise initially n-alkanes to the corresponding hydroperoxide and later reduce the intermediate by NADH, H^+ to alcohol (Ji et al., 2013). In addition to alcohol, some microorganisms convert the hydroperoxide intermediate

into the corresponding aldehyde. For instance, *Acinetobacter* sp. M-l is able grow rapidly on high molecular weight alkanes (Cl3 to C44) through oxidation via a n-alkane dioxygenase (Sakai et al., 1996).



(Reaction 2.3)

Methanotrophic bacteria oxidize CH_4 at the expense of O_2, producing CO_2 molecules that are excreted from the cells to the environment (Kalyuzhnaya et al., 2013). Methane monooxygenases (MMOs) are three-component enzymes, which are classified into membrane-associated particulate copper-dependent (pMMO) and soluble non-heme di-iron monooxygenases (sMMO) based on their solubility, sensitivity to cyanide and dependency on copper, (Kotani et al., 2006). Due to the intensive oxidizing potential of the active site of the enzyme, in addition to methane, this enzyme is able to catabolize several different (linear and cyclic) substrates (Murrell and Smith, 2010). The reductive potential of NAD(P)H assists the enzyme to split the covalent bond in O_2, incorporating it into methane structure to produce methanol. The other oxygen atom is combined with two hydrogen atoms to make a H_2O molecule (Murrell and Smith, 2010).

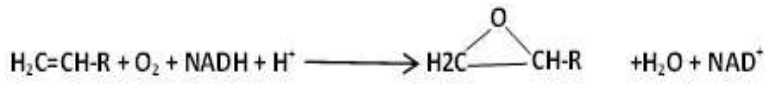


It is an important fact that most of the alkane degrading enzymes have a wide substrate range. This wide spectrum of substrate degradation ability is not just restricted to the common alkane monooxygenases as other monooxygenases, including the cytochrome containing monooxygenases and dioxygenases are able to degrade a large variety of substrates, too (van Beilen et al., 2006, Van Beilen et al., 2003). Furthermore, each microorganism may express more than one type of monooxygenases. *Pseudomonas putida* GPo1, for instance, produces a complex of a non-heme membrane associated monooxygenase (alkB), a soluble rubredoxin protein (alkG) and a rubredoxin reductase (alkT) that can catalize the initial step for the degradation of C5 to C12 *n*-alkanes (Whyte et al., 2002b).

2.1.4 Metabolism of alkenes

Unsaturated hydrocarbons (alkenes) are more sensitive than alkanes and aromatic hydrocarbons to degradation due to the higher reactivity of their double bonds (Grossi et al., 2011). Biodegradation of unsaturated aliphatic hydrocarbons (C2-C6) has been seen in a large variety of bacterial divisions (Beller et al., 2010, Wiegel, 2006, Grossi et al., 2008, Zhang et al., 2013c). In the same way as for alkane monooxygenase, several microorganisms can insert an oxygen atom into an alkene substrate via a monooxygenase activity, producing an epoxide molecule (Fig. 2.1.2) (Lin et al., 2011).

Epoxidation is a process in which O_2-/NADH-dependent monooxygenases, namely alkene monooxygenases (AMOs), add an oxygen atom to the double bonds of the unsaturated aliphatic hydrocarbons (Reaction 2.4) (Krishnakumar, 2007). Biochemically, based on a hypothesis by Weijers and colleagues (Weijers et al., 1995), a nucleophilic attack to the C1 or C2 of epoxides (1,2-epoxyalkanes and 2,3-epoxyalkanes, respectively) by the sulfhydryl group of an enzyme opens the epoxide ring and yields a β-hydroxythioether intermediate. Then, transfer of hydrogen from C-2 to NAD^+ converts the β-hydroxythioether to its corresponding β-ketothioether, which further undergoes a carboxylation attack to the terminal carbon by an epoxide carboxylase.



(Reaction 2.4)

AMOs, like MMO, are multi-component enzymes, three in *R. rhodochrous* B-276 and four in *Xanthobacter* Py2, belonging to the diiron oxygenase family with a highly conserved (binuclear or four nuclear, respectively)-iron core that interacts with Glutamate and Histidine residues (Krishnakumar, 2007, Smith, 2010). This core binds the O_2 molecule necessary for the hydroxylation and epoxidation of a broad range of unsaturated alkenes as well as a variety of chlorinated alkenes, such as tri-chloro-ethene, vinyl chloride and 1,3-dichloropropene (Krishnakumar, 2007, Taylor et al., 2010).

Some other alkene-utilizing bacteria produce flavin active-centre monooxygenases which are used especially for the production of chiral epoxides, particularly those compounds with aromatic substituents (Smith, 2010). *Rhodococcus* sp. ST-5 and ST-10 and several

Pseudomonas strains use a styrene monooxygenase (SMO) to epoxide the vinyl side-chain of styrene. This system is a two-component FAD-associated enzyme comprising of a flavin-oxidoreductase unit (SMOB) and a monooxygenase unit (SMOA) (Toda et al., 2012). In addition, some bacteria use cytochrome P450 system to insert an oxygen atom, obtained from molecular oxygen) into alkenes to form epoxides (Jin et al., 2003, Guengerich and Munro, 2013). *Olaromonas* sp. strain JS666 for instance, uses cytochrome P450 monooxygenase to trigger the oxidation of *cis*-1,2-dichoroethene (*c*DCE) in aerobic conditions (Nishino et al., 2013).

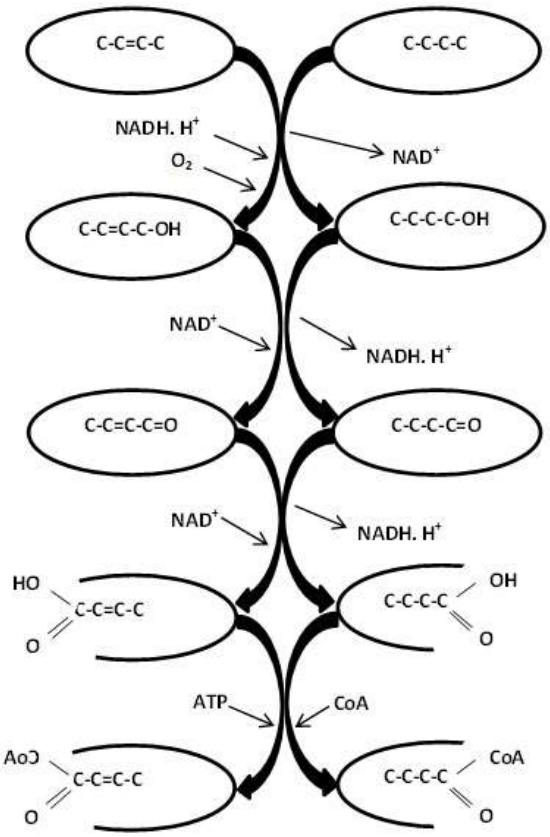
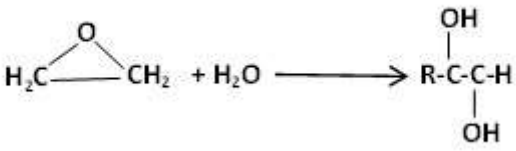


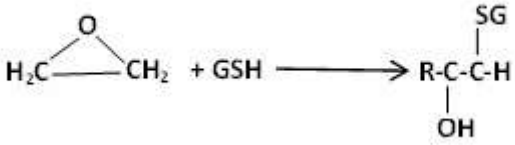
Figure 2.1.2 Biodegradation of butane and butene. This figure shows that both saturated and un-saturated C4 molecules are activated by a same pathway.

Although the broad substrate activity of alkene degrading bacteria is beneficial for remediation purposes, there are some limitations for their application. Sometimes the products of these enzymes are xenobiotic for themselves, requiring other enzymes to mineralize these resulting compounds (Jazestani, 2012). A bioremediation process will be most promising if the substrates mineralise completely to simple molecules such as CO_2 and H_2O (Megharaj et al.,

2011, Chen et al., 2014). Furthermore, in some cases, the product of these enzymes, especially epoxides, are toxic for the enzyme itself and other cellular proteins (Suttinun et al., 2013). These products are able to inactivate irreversibly the enzyme involved in the initial reaction of the biodegradation process and to interact with the other biological contents of cells (such as DNA, lipids and proteins, leading to their inactivation (Suttinun et al., 2013, KUROKI, 2012). Oxidation of trichloroethylene (TCE) by an alkane monooxygenase, produced by *Rhodococcus corallines*, yields a trichloroethylene epoxide, which is able to bind covalently both to the enzyme and to other cellular proteins (Suttinun et al., 2013). Therefore, those hydrocarbon degrading organisms that utilise epoxidation mechanisms have to neutralize these toxic compounds. In a manner similar to mammalian systems, several microorganisms use an oxide hydrolases or glutathione-S-transferases (GSTs) to neutralize epoxides via the addition of a water molecule or a glutathione originated hydrogen, respectively (Reactions 2.5 and 2.6) (Widersten et al., 2010, Wood, 2008).



(Reaction 2.5)



(Reaction 2.6)

In a completely different pathway, some bacteria use carboxylases for CO_2 assimilation through degradation of the epoxides, yielding corresponding β-keto acids. Growth of *Xanthobacter* Py2 in a media containing propene as the sole source of carbon, for instance, leads to production of epoxypropane, which is subsequently converted to acetoacetate via a CoM-dependent pathway (Glueck et al., 2010). In this pathway the resulting (S or R) epoxypropane is activated by a CoA and after assimilation of one CO_2 molecule through two reactions catabolized by a NAD dependent dehydrogenase and a NADP-dependent oxidoreductase is converted first to 2-ketopropyl-CoM and finally to acetoacetate, respectively (Reaction 2.7) (Erb, 2011).

(Reaction 2.7)

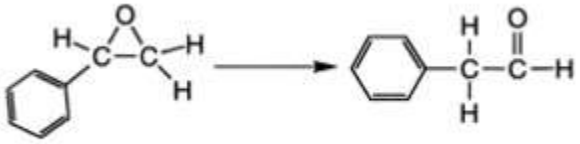
Propene + NADH, H^+ + O_2 → (R or S)epoxypropane + NAD^+

(R or S)epoxypropane + CoM → 2-(R or S)epoxypropyl-CoM

2-(R or S)epoxypropyl-CoM + NAD^+ → 2-Ketopropyl-CoM + $NADH,H^+$

2-Ketopropyl-CoM + CO_2 + NADPH → Acetoacetate + NADP

Since this reaction is regulated by CO_2, in the absence of this key molecule the catabolising bacteria use another strategy in which isomerases convert the epoxides to nontoxic compounds such as ketones and aldehydes (Glueck et al., 2010). In the case of the degradation of styrene by *Rhodococcus opacus* 1CP, for instance, Styrene oxide isomerase (SOI) converts the emerged styrene oxide to phenyl-acetaldehyde (Reaction 2.8) (Oelschlägel et al., 2012). Furthermore, growth of *Xanthobacter* strain Py2 in a media containing butylene-1,2-oxide or propylene oxide leads to production of methylethyl-ketone and acetone, respectively (Reaction 2.9) (Weijers et al., 1995). Also, spontaneous and slow degradation of the produced epoxides can sometimes yield their corresponding diol (Hou, 2006).



(Reaction 2.8)

(Reaction 2.9)

2.1.5 Branched chain alkanes

Biodegradation of branched chain alkanes, especially l3-branched (anteiso-) hydrocarbons, quaternary compounds and isoterpenoids is more difficult than for simple alkanes and alkenes, and most bio-degraders resort to complete degradation of these molecules (Rocha et al., 2011, Johnson et al., 2012). Although most investigators reported the branched chain alkanes to be very recalcitrant compounds, microorganisms are able to degrade these compounds (Alvarez et al., 2009).

Many microorganisms express an inducible enzymatic system that hydroxylates the terminals of different kinds of branched hydrocarbons to produce their corresponding acids or ketones

(Alvarez et al., 2009, van Hylckama Vlieg et al., 2000, Nhi-Cong et al., 2009). For instance, isoprene (2-methyl-1,3-butadiene) is initially oxidised by a monooxygenase expressed by isoprene degrading bacteria, such as *Rhodococcus* AD45, in which the methyl-substituted double bond is epoxidised to produce 1,2-epoxy-2-methyl-3-butene (Shennan, 2006, Alvarez et al., 2009). The epoxide ring of this intermediate molecule is opened by glutathione-*S*-transferase to produce 1-hydroxy-2-glutathionyl-2-methyl-3-butene (HGMB). Further oxidation of the terminal hydroxyl group by dehydrogenases yield an organic acid referred to as 2-glutathionyl-2-methyl-3-butenoic acid (GMBA) (Shennan, 2006, van Hylckama Vlieg et al., 2000).

Many highly branched hydrocarbons, such as pristane (2,6,10,14-tetramethylpentadecane), farnesane (2,6,1O-trimethyldodecane), norpristane (2,6,1O-trimethylpentadecane) and phytane (2,6,10,14-Tetramethylhexadecane) are initially activated by a hydroxylation mechanism (Thi Nhi Cong, 2008). Pristane is a highly persistent isoprenoid alkane found in microorganisms and also higher organisms, which is derived from the degradation of tocopherols at high temperatures and also from the breakdown of methyl-trimethyl-tridecylchromans (Schurig et al., 2014). This compound is sensitive to enzymatic degradation by many microorganisms such as *Rhodococcus ruber* and *Mycobacterium neoaurum*, and is degraded through either or both of ω-hydroxic acids and α-ω-dioic acids, yielding terminal mono or di-carboxylicic acids (Fig. 2.1.3) (Nhi-Cong et al., 2009).

2.1.6 Cycloaliphatic compounds

Several types of cycloaliphatic hydrocarbons such as cyclopentane, methylcyclopentane and cyclohexane are found extensively in crude oils and some petroleum products (Comandini et al., 2013). In addition to these oil originated compounds, alicyclic hydrocarbons are released into the environment via different natural activities, including through microbial or plant production of terpenoids and lipids (Wilbon et al., 2013, Gupta et al., 2014), and artificial manmade activities, such as production of nylons, flavours, herbicides, insecticides and solvents (Wittcoff et al., 2012). Although these compounds are hardly degraded by microbial activities, due to the low solubility of these hydrocarbons and their toxic effects on biological systems, many bacteria are able to metabolise these hydrocarbons through an oxidative pathway, converting these compounds to adipic acid (Iwaki et al., 2008, Comandini et al., 2013).

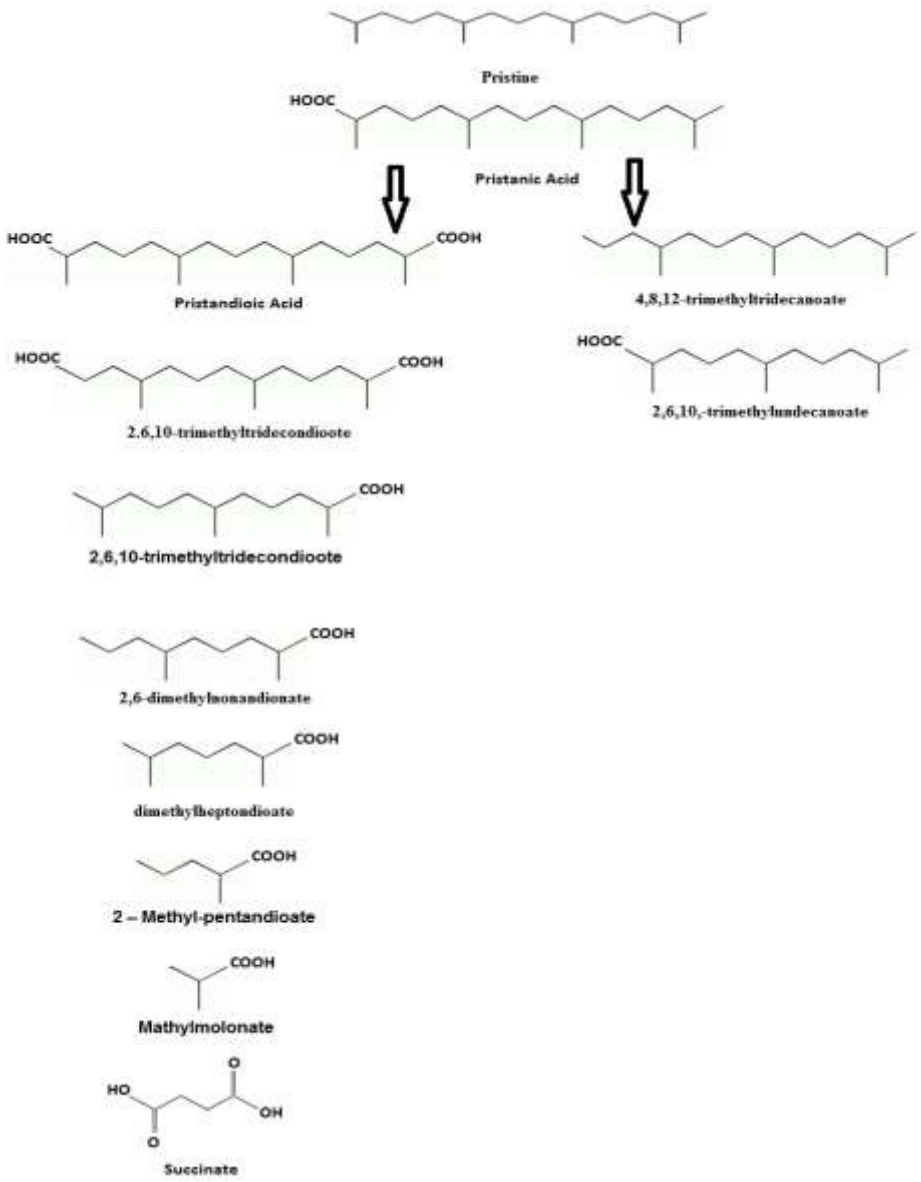


Figure 2.1.3. Biodegradation of pristane. The resulting pristanic acid can pass through two different pathways. Pristandioic acid is degraded to succinate, which is one of critical intermediates in microbial metabolism. The fate of 2,6,10-trimethylundecanoate is not clear yet.

Biochemical and genetic studies on the degradation of cyclohexane by *Acinetobacter* sp. showed involvement of five enzymes encoded by two operons (chn BER ORF and chnADC ORF). These two operons are located in the chromosome in an opposite direction to each other and are regulated by a regulatory protein encoded by *chn*R). After hydroxylation of

cyclohexane by cyclo-hexanone monooxygenase (encoded by *chn*B) at the expense of NAD(P)H and O_2, the produced cyclohexanol is converted by NAD(P)H-dependent aldehyde dehydrogenase (encoded by chnE) to cyclohexanone in which a NAD(P) is reduced to NAD(P)H (Fig. 2.1.4). Incorporation of one atom of oxygen into the ring of the resulting cyclohexanone by cyclohexanone monooxygenase (encoded by chnA) through oxidation of NADPH and reduction of O_2 to H_2O converts this intermediate to caprolactone. After breaking the ring through insertion of one molecule of H_2O into the ring via caprolactone hydrolase (encoded by chn C), the produced 6-hydroxy hexanoic acid is further oxidized by 6-hydroxyhexanoic acid dehydrogenase (encoded by chnD) to 6-oxohexanoic acid and then by NAD(P)+-dependent aldehyde dehydrogenases to adipic acid (Koma et al., 2004, Comandini et al., 2013).

Rhodococcus ruber SC1 oxidizes cyclododecane as its sole source of energy and carbon in a way analogous to the biodegradation of cyclohexane (Fig. 2.1.5) (Kostichka et al., 2001b). A monooxygenase initiates oxidation of the compound in which the bacterium converts cyclododecane to cyclododecanol at the expense of one NAD(P)H and one O_2. A similar process involving a NAD(P)-dependent cyclododecanol dehydrogenase oxidizes the resulting cyclododecanol into cyclododecanone. From this point, the products of a cluster of four genes (*cddABCD)* promote the oxidation process of cyclododecanone: a monooxygenase (encode by *cdd*A) inserts an atom oxygen into the ring and converts the emerged cyclododecanone to lauryl lactone, which in turn accepts one molecule of H_2O as a result of the action of lactone hydrolase (encoded by *cdd*B) to produce an acidic intermediate (12-hydroxy lauric acid). Two NAD(P) dependent dehydrogenases, namely 12-hydroxylauric acid dehydrogenase (encoded by *cdd*C) and 12-oxolauric acid dehydrogenase (encoded by *cdd*D) convert this intermediate to 12-oxolauric acid and finally to DDDA (dodecanedioic acid) (Chen et al., 2004).

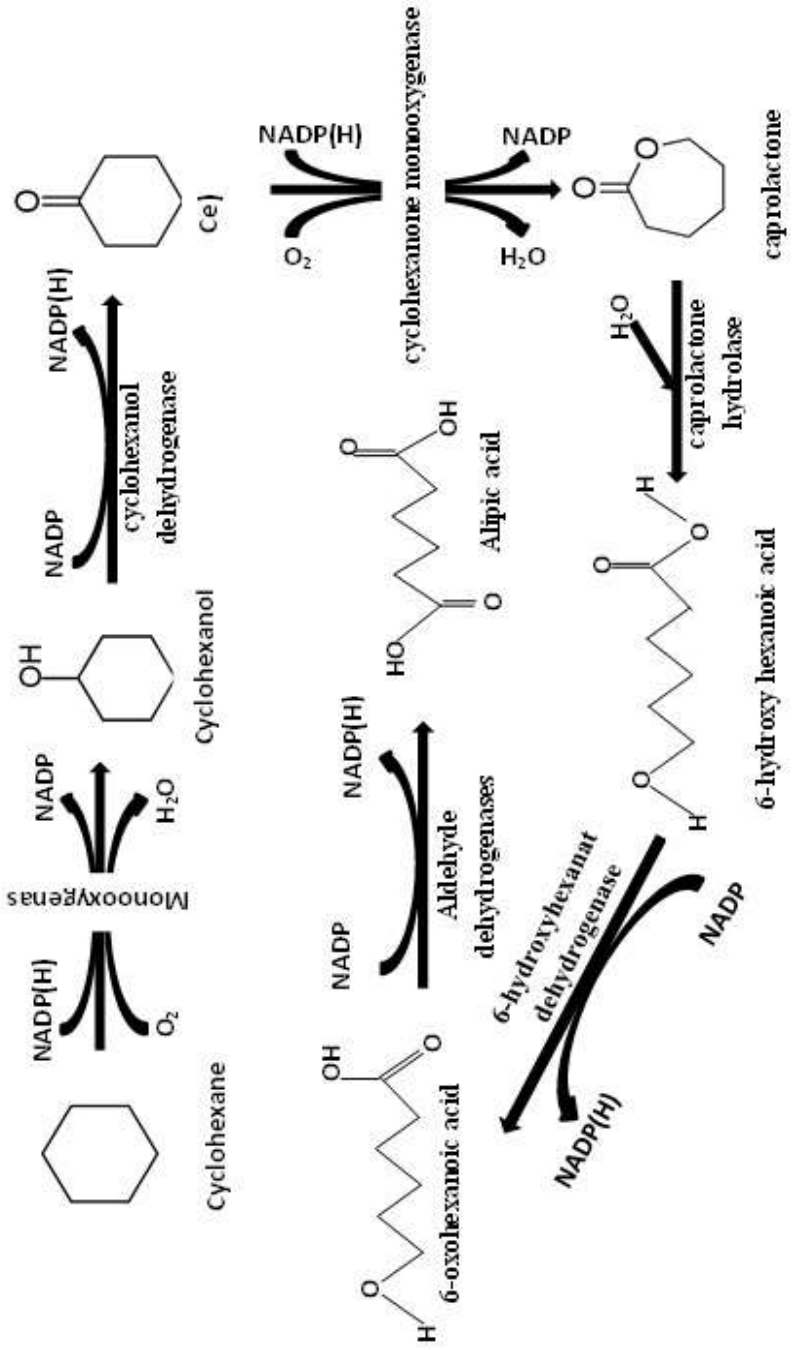


Figure 2.1.4. Biodegradation of cyclohexane. The molecule is linearized through insertion of one atom of oxygen into the ring. The emerged 6-hydroxy hexanoic acid is oxidized through two NADP dependent oxidation to Alipic Acid.

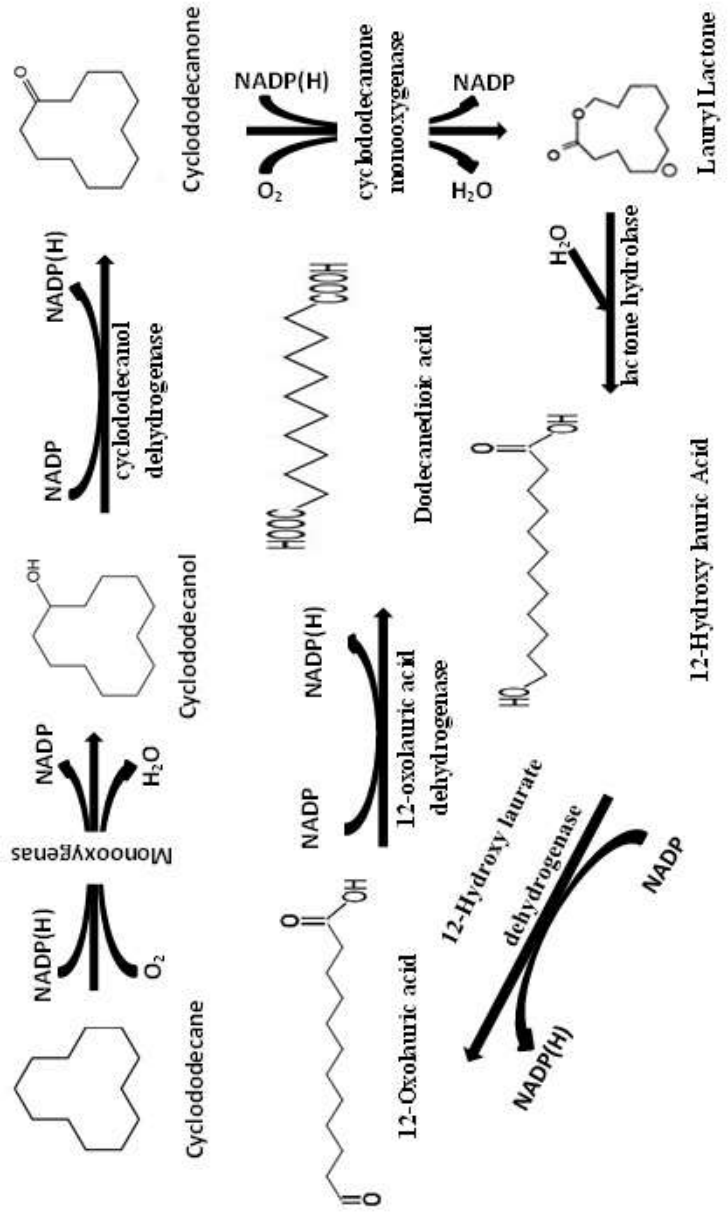


Figure 2.1.5. Conversion of a cyclododecane to dodecandioic acid (DDDA) in *Rhodococcus* ruber CD4

Although several microorganisms are able to independently degrade cycloaliphatic compounds to simpler molecules, these compounds can also be catabolised by a mixed community of microorganisms through a mechanism referred to as co-metabolism (Lee and Cho, 2008, Morgan and Watkinson, 1994). A consortium of *Rhodococcus* spp., *Sphingomonas* sp. and

Stenotrophomonas sp., for instance, use cyclohexane as their sole source of energy and carbon (Lee and Cho, 2008). Also, a consortium of *Rhodococcus* sp. (80 to 90% of the population), *Pseudomonas cepacia* and *Flavobacterium* sp. (together, 10-20% of the population) has been reported with the ability to completely degrade cyclohexane to CO_2 and H_2O. In this last consortium hydrocarbons are initially oxidised by a monooxygenase to alcohols or ketones, which are then consumed by other members of the community (Morgan and Watkinson, 1994).

2.1.7 Anaerobic hydrocarbon biodegradation

Microbial degradation of many different substrates, especially recalcitrant hydrocarbons, is very limited under anaerobic conditions, due to a requirement for molecular oxygen as the final electron acceptor, slowing down these biological activities (Widdel et al., 2010b). However, based on a large variety of studies on microorganisms and the G_0 values of the anaerobic degradation reactions, anaerobic metabolism of hydrocarbons by microorganisms is energetically possible (Rosenberg et al., 2007, Jobelius et al., 2011). Slow degradation of these molecules by anaerobic bacteria provides the required energy and carbon sources for the growth of other catabolizing microorganisms and methane that is used by many other aerobic organisms, including some animals, that are dependent on methane oxidising microorganisms (Michaelis et al., 2002). A large diversity of facultative and obligatory anaerobic microorganisms (bacteria and archaea) have been isolated with the ability to anaerobically degrade hydrocarbon molecules. These microorganisms utilize anaerobic respiration (using nitrate, nitrite, nitrous oxide, sulphate, thiosulfate, carbonate and metal ions), or fermentation or anoxygenic phototrophic reactions, for the donation of electrons and hydrogen for substrate catabolism activities (Grossi et al., 2008). For anaerobic microorganisms, both sulphate and nitrate reducing bacteria, shorter chain length alkanes are more recalcitrant than mid- to long-chain length alkanes (Hasinger et al., 2012b). Due to the lack of evaporation of short length hydrocarbons (up to n-C17) in anaerobic conditions, these compounds can build up and exert a toxic effect on the cell membranes of microorganisms, inhibiting their degradation (Hasinger et al., 2012b). Furthermore, for sulphate reducing bacteria degradation of branched alkanes, such as pristane and phytane, is more efficient than for normal alkanes (Hasinger et al., 2012b).

Anaerobic hydrocarbon decomposition is usually found in deep and anoxic conditions such as natural oil and gas leakages/seeps on land or the sea floors and those sites polluted with oil or its by-products (Jobelius et al., 2011). Furthermore, this type of biodegradation can occur

below the surface of areas exposed to oxygen where aerobic biological activity ceases as all the oxygen is utilised (Singh et al., 2014a). In most cases, degrading activity of aerobic microorganisms in a low gradient of oxygen limits oxygen availability, leading to a succession of anaerobic hydrocarbon degrading microorganisms (Singh et al., 2014a). After oxygen exhaustion, there is a sequential employment of the electron acceptors (nitrate, ferric iron, sulphate and H_2 subsequently) to supply the highest potential energy from the degradation of hydrocarbons (Dojka et al., 1998, Pilloni et al., 2011). Indeed, in this microbial succession the changes are based on the available energy and electron acceptor systems, beginning with nitrate and nitrite reducing microorganisms, followed by metal ($Mn^{+}5$ or $Fe^{+}3$, U^{+6}) reducing bacteria, sulphate-reducing bacteria and finally methanogenic microorganisms (Pilloni et al., 2011).

2.1.7.1 Denitrifying bacteria. These hydrocarbon degrading bacteria belong to many different microbial phyla, mainly to the Betaproteobacteria (order Rhodocyclales and genera *Azoarcus, Georgfuchsia*, *Thauera* and *Aromatoleum*), *Castellaniella defragrans* (a species belonging to the family Alcaligenaceae and the order Burkholderiales) and *Magnetospirillum* (belonging to the Alphaproteobacteria) (Brodkorb et al., 2010b, Shinoda et al., 2004, Weelink et al., 2009), and degrade different aliphatic and aromatic hydrocarbons at the expense of nitrate reduction (Reaction 2.10) (Weelink et al., 2009).

(Reaction 2.10) $$\mathbf{5\ Hexadecane\ (C_{16}H_{34}) + 98\ NO_3^- + 18H^+ \rightarrow 80\ HCO_3^- + 49\ N_2 + 54\ H_2O}$$

2.1.7.2 Sulphate-reducing microorganisms. Many of these bacteria, including Deltaproteobacteria (like *Desulfobacula toluolica, Desulfococcus oleovorans, Desulfatibacillum alkenivorans, Desulfosarcina cetonicum, Desulfoglaeba alkenexedens, Desulfatibacillum aliphaticivorans*) (Callaghan et al., 2012b), *Desulfotomaculum* (Selesi and Meckenstock, 2009b) and *Archaeoglobus fulgidus* (an archeon) (Khelifi et al., 2010) are able to degrade long/short chain saturated/unsaturated aliphatic hydrocarbons, non-substituted aromatics (such as naphthalene and benzene) and alkyl-benzenes (such as like xylenes, toluene and ethylbenzene). *Desulfatiferula* strains BE2801 and LM2801T (Grossi et al., 2008) and *D. aliphaticivorans* strain $CV2803_T$ (Grossi et al., 2007) Sulfate reducing bacteria oxidise the double bond of 1-alkenes, converting them to the corresponding 1-alkanols and further to fatty acids and finally to CO_2 (Reaction 2.11). Hydratase, a reversible enzyme with the ability to

catalyse the rearrangement of hydrogen and hydroxyl groups on double bonds, plays a crucial task in the production of secondary alkanols (substitution of -Ol on C2-C5) from 1-alkenes (Rontani et al., 2002). In addition to the oxidative attack, *D. aliphaticivorans* strain CV2803$_T$ (Grossi et al., 2007) is able to add an organic residue (such as methyl or ethyl) to alkanes and alkenes, yielding the corresponding methyl/ethyl- fatty acids (Grossi et al., 2011).

(Reaction 2.11) $$\mathbf{4\ Hexadecane\ (C_{16}H_{34}) + 49\ SO_4^{2-} + 34\ H^+ \rightarrow 64\ HCOO^{3-} + 49\ H_2S + 4H_2O}$$

2.1.7.3 Metal oxidizing microorganisms. In addition to nitrogen and sulphur containing inorganic compounds, some bacteria are able to use other inorganic compounds as electron recipients while metabolising hydrocarbons. For instance, *Geobacter* and *Georgfuchsia* reduce metal-ions (Fe(III), Mn(IV) or U(VI)) through anaerobic degradation of hydrocarbons (Zengler et al., 1999, Weelink et al., 2009). In addition to these genera, a toluene degrading bacterium (strain G5G6), belonging to the Betaproteobacteria is able to mineralize toluene completely to CO_2 while utilizing Fe^{3+}, Mn^{4+} and NO^{3-} as its terminal electron acceptors (Weelink et al., 2009). Furthermore, *Dechloromonas aromatica* RCB (Coates et al., 2001) and strain Y5 (Liu et al., 2004) oxidize toluene in reactions coupled to (per)chlorate and arsenate, respectively. Since sites contaminated with oil products may contain arsenic in addition to aromatic compounds, this ability of microorganisms to degrade aromatics in association to As(V) reduction is beneficiary in terms of environmental remediation (Liu et al., 2004).

2.1.7.4 Intra-aerobic anaerobes. This category of microorganisms are able to extract oxygen from chlorate, perchlorate, nitric oxide (NO) or nitrate in order for it to be utilized as an electron acceptor (Wu et al., 2011, Ettwig et al., 2012). *Alicycliphilus denitrificans* and *Pseudomonas chloritidismutans*, for instance, possess two periplasmic molybdenum-containing enzymes, chlorate reductase and perchlorate reductase, which enable them to reduce chlorate and perchlorate to chlorite. Chlorite, in turn, is immediately reduced by a heme-containing enzyme (chlorite dismutase) into Cl_2 and O_2 (Reaction 2.12) (Wardlaw et al., 2008, Ettwig et al., 2012).

(Reaction 2.12) $$\mathbf{ClO^{2-} \rightarrow Cl^- + O_2\ \ (\Delta G^{o/} = -100\ kJ\ mol^{-1}\ O_2)}$$

Furthermore, several intra-aerobic-anaerobe bacteria are able to employ a denitrification pathway for obtaining molecular oxygen by breaking nitric oxide or nitrate down to N_2 and O_2 (Reaction 2.13). The extraction of O_2 by denitrifying bacteria emerged through investigations on the oxidation of methane by a pure culture of a methane oxidizing microorganism (*Methylomirabilis Oxyfera*) in which the catabolism of methane to CO_2 is associated with nitrate reduction (Reaction 2.13 and 2.14) (Smith et al., 2011).

(Reaction 2.13) $\mathbf{2NO_2^- \rightarrow N_2 + O_2 \quad (\Delta G^{0/} = -173 kJ\ mol{-}1\ O_2)}$

(Reaction 2.14) $\mathbf{3\ CH_4 + 8\ NO^{-2} + 8\ H^+ \rightarrow 3\ CO_2 + 4\ N_2 + 10\ H_2O \quad (\Delta G^{o/} = -928\ kJ\ mol{-}1\ CH4)}$

2.1.7.5 Methanogenesis. The production of methane and CO_2 from the anaerobic degradation of squalene (a methyl-branched alkene with formula $C_{30}H_{50}$) and 1-hexadecene was first described by Schink (Schink, 1985). Alkenes are initially activated by addition of a H_2O to the double bond or by addition of a hydroxyl to the carbon atom adjacent to the unsaturated bond (Reaction 2.15) (Grossi et al., 2008). In the case of hexadecane, this compound is first converted to hexadecan-1-ol and after oxidation to hexadecanoic and palmitate, which are broken down to acetate units and are finally catabolised to methane (Reaction 16) (Morgan and Watkinson, 1994).

(Reaction 2.15) **Palmitate:** $\mathbf{C16H32 + 12\ H20 \rightarrow 12\ CH4 + 4HCO^{3-} + 4H^+}$

(Reaction 2.16) **Hexadecan:** $\mathbf{4(C_{16}H_{34}) + 45\ H_2O \rightarrow 49\ CH_4 + 15HCO^{3-} + 15H^+}$

Several anaerobic microorganisms are able to mineralize hydrocarbons to methane as part of microbial syntrophic consortia. Such a consortium can be seen in the syntrophic relationship between methanogenic archaea and hydrogen-producing bacteria for the degradation of aromatic hydrocarbons to methane and CO_2. In this cooperation, the H_2 generated by special hydrocarbon degrading microorganisms, such as *Smithella* sp., *Syntrophus* sp. and *Desulfotomaculum* sp. is used for reduction of CO_2 by methanogenic archaea to produce methane (Reactions 2.17 and 2.18) (Grabowski et al., 2005, Jones et al., 2008).

(Reaction 2.17) $\mathbf{Toluene + 14\ H_2O \rightarrow 7CO_2 + 18\ H_2}$

(Reaction 2.18) $\mathbf{4H_2 + CO_2 \rightarrow CH_4 + 2H_2O}$

2.1.7.6 Anoxygenic sulphate dependent phototrophic bacteria. In addition to the anaerobic respiratory microorganisms, some anoxygenic sulphate dependent phototrophic bacteria (such as *Blastochloris sulfoviridis)* (Zengler et al., 1999) and fermentative prototrophs (such as *Pelobacter acetylenicus* belonging to Deltaproteobacterium) (Einsle et al., 2005) have been isolated that are able to degrade hydrocarbons anaerobically. These bacteria comprise only a small population of a microbial community; for instance, toluene degrading phototrophs comprise only 1% of the cultivable lithotrophic microorganisms. Because phototrophic microorganisms obtain their energy from light, the whole organic substrate is probably converted into cell components ($C_4H_8O_2N$) (Reaction 2.19) (Zengler et al., 1999).

(Reaction 2.19) $\mathbf{(17C_7H_8 + 25HCO^{3-} + 36\ NH_4^+ \rightarrow 36C_4H_8O_2N + 3H_2O + 11H^+)}$

2.1.8. The Biochemistry of Anaerobic Hydrocarbon biodegradation

Anaerobic biodegradation pathways are more diverse than aerobic catabolism and can be divided into 5 groups, including addition of a fumarate to methylene or methyl groups of hydrocarbons (Selesi et al., 2009), Oxygen-independent hydroxylation on 2nd or 3rd terminal C-atoms (to make secondary or tertiary alcohols) (Kniemeyer and Heider, 2001a), Carboxylation of unsubstituted carbon atoms of aromatics (Meckenstock and Mouttaki, 2011), Hydration of the double and triple bond of alkenes and alkynes (Tenbrink et al., 2011) and Reverse methanogenesis (Thauer, 2011).

2.1.8.1 Fumarate Addition Reactions

A variety of anaerobic hydrocarbon degraders, such as sulphate reducing bacteria, denitrifying microorganisms, iron-reducing bacteria, methanogenic consortia and anaerobic phototrophic microorganisms add a carboxyl group to a hydrocarbon in order to achieve greater degradation (Verfürth et al., 2004, Beasley and Nanny, 2012, Tierney and Young, 2010). The target carbon atom for the addition of the carboxyl group is different depending on the microbial species; C1 in SRBs, C1, C2 and C3 in *Desulfatibacillum aliphaticivorans* (Grossi et al., 2007).

The anaerobic degradation of toluene via fumarate is an additional system that has been widely studied in SRBs and denitrifying bacteria, which is used as a prototype example of this kind of metabolic pathway (Fig. 2.1.6) (Leuthner and Heider, 2000a). First, the addition of a

fumarate by oxygen-sensitive benzylsuccinate synthase (BSS) enzyme converts toluene to (R)-benzylsuccinate. In this reaction, removing one atom of hydrogen from the methyl group of toluene converts the molecule to a benzyl radical intermediate. This radical is very unstable and attacks the unsaturated bond of a fumarate molecule regaining the hydrogen atom (associated with the enzyme) to form a benzylsuccinate. The benzylsuccinate is then substituted with succinyl-CoA in preparation for further oxidative processes. Throughout this reaction, a succinate is released, which is converted (by succinate dehydrogenase) to fumarate in order to regenerate the fumarate source. Succinyl-CoA functions as CoA-SH donor and is added to the complex via a CoA transferase enzyme (BbsEF) in order to activate benzylsuccinate. Finally, through some β-Oxidation processes, a benzyl-CoA is produced and the succinyl-CoA is released for continuing the cycle. Benzoyl-CoA is the critical intermediate in the anaerobic catabolism of aromatic hydrocarbons that is further oxidized via breaking the aromatic ring (by benzoyl-CoA reductase) thereby converting it to aliphatic compounds that in turn are attacked by further hydrolytic and oxidative reactions, such as β-oxidation. A similar reaction has been seen in the degradation of xylene, o-toluidine and the cresol isomers by SRBs (Rotaru et al., 2010, Leuthner and Heider, 2000a).

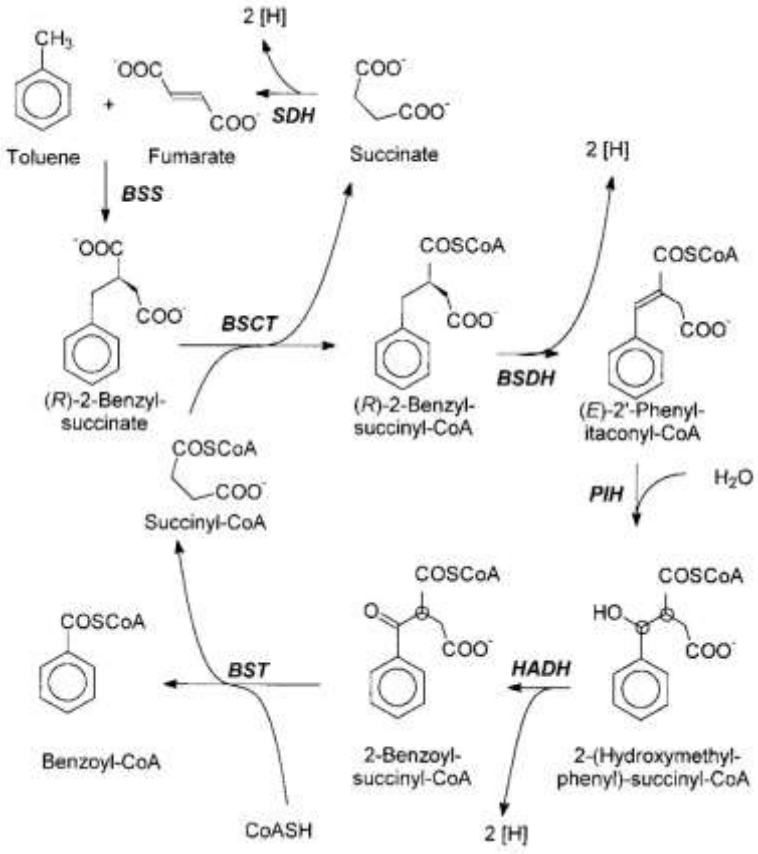


Figure 2.1.6. Catabolism of toluene via addition of fumarate to the methyl residue. In this route, toluene is activated to benzyl-CoA that is further oxidized to simplest metabolites. The

release succinyl-CoA is oxidized to succinate and further to fumarate to supply the fumarate pool. BSS: benzylsuccinate synthase; BS-CT: succinyl-CoA:(R)-benzylsuccinate CoA-transferase; SDH: Succinate dehydrogenase**;** BS-DH benzyl succinyl-dehydrogenase; PIH: (*E*)-Phenylitaconyl-CoA hydratase; HADH, 3-hydroxyacyl-CoA dehydrogenase **;** BST, benzoylsuccinyl-CoA thiolase;

Also, many SRBs and denitrifying bacteria are able to use this system for the degradation of a broad spectra of alkanes and cycloalkanes, from three to twenty carbons, depending on the bacterial species (Fig. 2.1.7) (Kropp et al., 2000, Cravo-Laureau et al., 2005). This reaction is initiated by substitution of a fumarate to the sub-terminal methylene carbon of the alkanes, and thereby converts the straight chain hydrocarbon to a branched compound (1-methylalkyl-succinate). After substitution of a CoA-SH to the terminal methyl group of the branched methylalkyl-succinate by succinate-CoA, the terminal methyl CoA is transferred to one carbon before the branch to generate a 4-methyl-branched fatty-acyl-CoA molecule. This intermediate is easily catabolized to one propionyl-CoA and acetyl-CoA units via the β-oxidation reaction. The produced propionyl-CoA is used to recharge the fumarate supply of the pathway (Rabus et al., 2001, Wilkes et al., 2002).

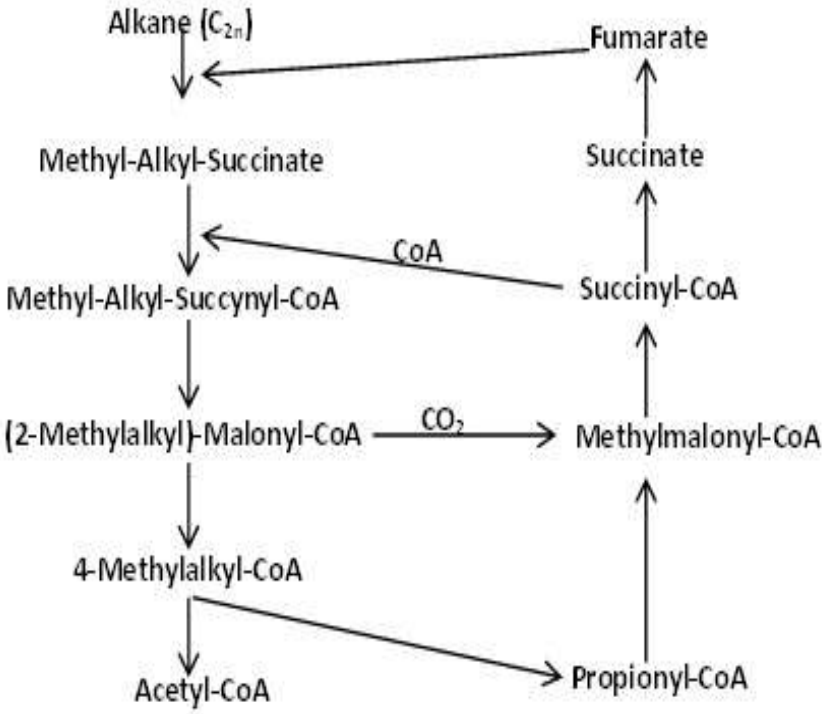


Figure 2.1.7. Addition of fumarate to alkanes. The emerged propionyl-CoA is converted to fumarate in order to continue this reaction.

The similarities between the small subunit of the enzyme involved in the addition of the carbon group to the alkane with the corresponding enzyme involved in the anaerobic activation of toluene indicated a similar process for the activation of aliphatic and aromatic compounds in anaerobic conditions (Widdel and Rabus, 2001). Furthermore, based on sequencing studies, the genes involved in this process for both denitrifying bacteria and SRB (ass-1 and ass-2 operons) are the same as for other fumarate-adding enzymes (Callaghan et al., 2012b).

2.1.8.2 Oxygen-Independent Hydroxylation

Although ethylbenzene is mainly degraded through the fumarate addition pathway, several denitrifying bacteria use an anoxic hydroxylation reaction in which a periplasmic ethylbenzene dehydrogenase (EBDH) hydroxylates the terminal carbon of the side chain of this molecule to produce S-1-phenylethanol (Fig. 2.1.8) (Heider, 2007). This enzyme is a soluble periplasmic molybdenum-containing enzyme belonging to the DMSO reductase family (Heider, 2007) with three different subunits (α, β and γ). Two of the subunits (αand β) are homologous to the *Escherichia coli* nitrate reductase) in terms of protein structure and amino acid sequences (Kloer et al., 2006). In addition, the active site of this enzyme contains five iron-sulphur compounds (αand β in subunits) and a heme b cofactor (placed in γ subunit)(Kloer et al., 2006). After the movement of the resulting (S)-1- phenylethanol through the cell membrane via passive diffusion (Kniemeyer and Heider, 2001b), the molecule is firstly oxidized to acetophenone (by NAD-dependent alcohol dehydrogenase) and then the product is carboxylated by acetophenone carboxylase (APC) to benzoyl-acetate (Rabus et al., 2005, Jobst et al., 2010). At the next step, a thioesteration process by benzoylacetate-CoA ligase (BAL) transfers a CoA to the molecule and the benzoyl-acetyl-CoA is finally is cleaved to benzoyl-CoA and acetyl-CoA units. *Aromatoleum aromaticum* PbN1 is a denitrifying bacterium with an ability to catabolize n-propylbenzene through the same biochemical shunt as is used in the degradation of ethylbenzene, leading to the production of benzoyl-CoA and propionyl-CoA (Kniemeyer and Heider, 2001b).

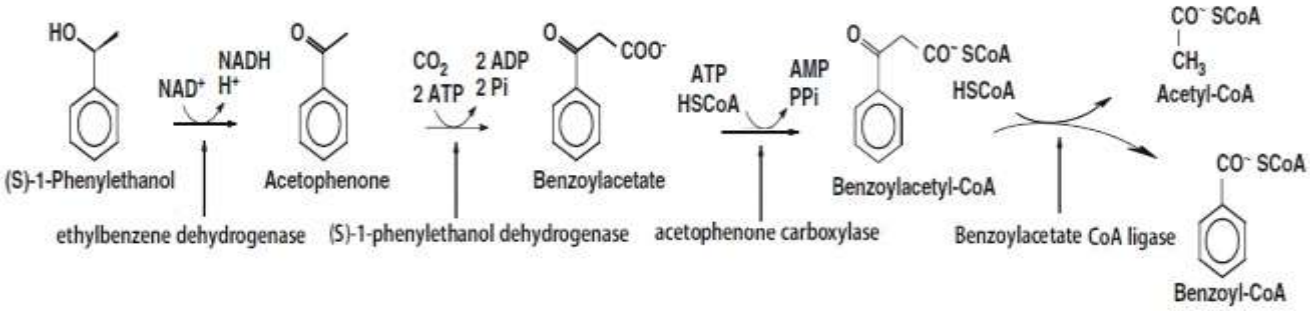


Figure 2.1.8. Oxygen-Independent hydroxylation of phenylethanol. The first enzyme (ethylbenzene dehydrogenase) is found in the periplasmic space while other enzymes are present in the cytosol.

2.1.8.3 Carboxylation

Many anaerobic bacteria use carboxylase enzymes at the first step of anaerobic hydrocarbon degradation to add a carboxyl group directly to both aliphatic, such as propylene, and aromatic hydrocarbons, such as benzene, phenanthrene, naphthalene and biphenyl, converting them to their corresponding carboxylate forms (Abu Laban et al., 2010, Jobst et al., 2010). Growth of *Xanthobacter* strain Py2 on propylene oxide is a clear example in which the bacterium assimilates CO_2 to utilize the substrate. In this reaction, *Xanthobacter* first isomerises the propylene to an enol or keto form of acetone and then binds the assimilated CO_2 to the acetone to produce acetoacetate, which in turn is broken down to two acetates (Small and Ensign, 1995). Also, growth of *R. rhodochrous* B-276 on 2-propanol and acetone induce expression of an acetone carboxylase that adds an inorganic CO2 to acetone to yield an acetoacetate (Clark and Ensign, 1999).

2.1.8.4 Unsaturated bond (Alkene and Alkyne) Hydration

Anaerobic degradation of alkene via a hydration reaction is a recently identified reaction in which microorganisms add a H_2O molecule to the hydrocarbon, converting it to a primary alcohol (Grossi et al., 2007). This mechanism was firstly found in *Castellaniella defragrans* in which the bacterium produces a periplasmic enzyme, namely linalool dehydratase/isomerase (LDI), capable of catalysing reversibly the hydration of linalool hydrocarbon to an unsaturated hydrocarbon (myrcene) and of isomerization of myrcene to geraniol (Fig. 2.1.9). This geraniol is then transferred to the cytoplasm where it is oxidized to granate (Brodkorb et al., 2010a).

Some bacteria such as *N. rhodochrous* and *Pelobacter acetylenicus* are able to use hydrating reactions to mineralize several unsaturated alkenes and alkynes such as ethine and acetylene (Rosner and Schink, 1995, Widdel et al., 2010b). This bacterium adds a water molecule to acetylene via an enzymatic reaction catalysed by acetylene hydratase and produces a C_2-enol intermediate that is immediately converts to acetaldehyde (Reaction 2.20). Acetylene hydratase is an enzyme containing molybdenum, tungsten, Fe_4S_4 as well as guanine dinucleotide, belonging to the DMSO-reductase family (Tenbrink et al., 2011). The acetaldehyde is further oxidised to acetate, which is activated by CoA to be consumed in cellular metabolism (Rosner and Schink, 1995, Widdel et al., 2010b).

(Reaction 2.20) $C_2H_2 + H_2O \rightarrow CH_3CHO$

$2\ CH_3CHO + 3H_2O \rightarrow H_3C–COO^- + C_2H_5OH + 5\ H^+$

$H_3C–COO^- + ATP + CoA \rightarrow$ Acetyl-CoA

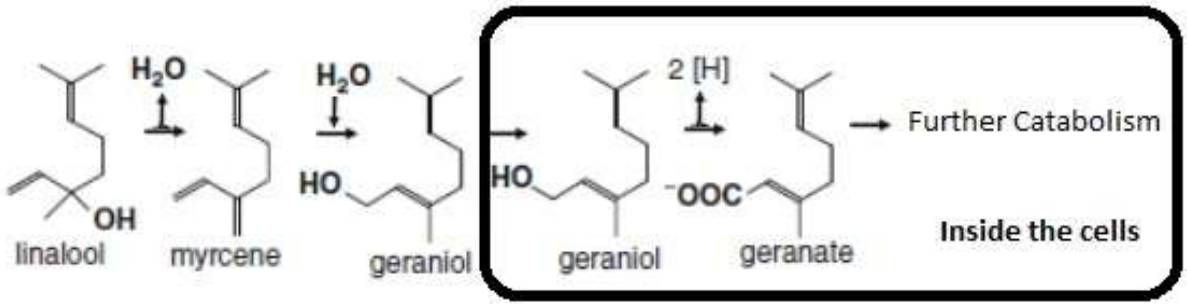


Figure 2.1.9. Hydration of Linalool. The molecule is first dehydrated and then re-hydrated in the extracellular environment. The resulting geraniol is transferred into the cell and is oxidized to geranate.

2.1.8.5 Reverse Methanogenesis

Anaerobic marine environments provide suitable conditions for a group of Mn(IV) and Fe(III) reducing bacteria that use a reverse methanogenic reaction as the initial step for the degradation of methane (Beal et al., 2009). The final methanogenic reaction in all methanogens is catalysed by methyl-coenzyme M reductase (Mcr) that binds to coenzyme B (CoBSH) via methyl-coenzyme M (CoMSH) to produce a complex of CoM-S-S-CoB-heterodisulfide and methane.

Several methanotrophic microorganisms use this enzyme for the initial activation of methane and take up the molecule in the form of Methyl-Co-M (Thauer, 2011).

In addition to independent microorganisms, several consortia of SRB containing specialised archaea belonging to the methanosarcinales are able to use the reverse methanogenesis process (Wilms et al., 2007). The H^+ and electrons produced through oxidation of CH_4 by methanotrophs are transported to the SRBs in order to reduce sulphate (Reaction 2.21). Although the mechanism in which the archaea are able to transfer the redox equivalents to their SRB partners is not fully understood, existence of an electron transport shuttle (Danko and Freedman, 2008) or some end-products such as methyl-sulphides (Moran et al., 2008), formate (Seeliger et al., 1998) and acetate (Brodersen et al., 1999, Aeckersberg et al., 1991) are hypothesised.

Reaction 2.21:

$$\mathbf{CH_4 + 3\ H_2O \rightarrow HCO_3^- + 9\ H^+ + 8\ e^-} \qquad E\ in\ situ = -0.289\ V\ (E^{o\prime} = -0.239\ V)$$

$$\mathbf{SO4^{2-} + 9\ H^+ + 8\ e^- \rightarrow HS^- + 4\ H_2O} \qquad E\ in\ situ = -0.248\ V\ (E^{0\prime} = -0.217\ V).$$

$$\textbf{Overall:} \quad \mathbf{CH_4 + SO4^{2-} \rightarrow HCO_3^- + HS^- + H_2O}$$

2.1.9 Application of hydrocarbon degrading enzymes in biotechnology

While some bacteria are killed directly by hydrocarbons and/or its products, some microbial populations are able to degrade these components as their nutritional source. These microorganisms may be commercially useful for applications in several different processes such as the production of special enzymes and chemicals and for cleaning up hydrocarbon contaminants (crude oil and/or its products) from different environments (Boopathy et al., 2012).

2.1.9.1 Chemicals and enzymes

Several microorganisms are used for the production of solvent chemicals from the degradation of hydrocarbons. *Arthrobacter* sp., for instance, is able to produce a large quantity of routine solvents through the degradation of naphthalene and *n*-hexadecane (Efroymson and Alexander, 1991). Also, methanotrophs have been abundantly used for the biotransformation of several organic chemicals (such as propylene to epoxypropane) and for the production of several biological compounds such as bioplastics, single cell protein (SCPs) and astaxanthin (Jiang et al., 2010).

Furthermore, some of the hydrocarbon degrading enzymes expressed by microorganisms are potentially of use in biotechnology. Naphthalene dioxygenase (NDO), for instance, is the initial enzyme for the degradation of naphthalene and is used in several different industries for the production of some medicines such as crixivan (a HIV protease inhibitor) (Seo et al., 2010), production of enantiopure arene *cis*-diols (prostaglandin E2 and polyphenylene) (Boyd et al., 2011) and the conversion of indoles into indoxyls and glucose to inositols (Zhang et al., 2013b). Furthermore, NDO is a crucial enzyme involved in the degradation of many aromatic hydrocarbons from soil and aquatic habitats and thus is considered as a useful enzyme in Biodegradation (Librando and Pappalardo, 2013).

Cytochrome P450 monooxygenase is an enzyme utilised in several industries for both bioremediation and the synthesis and improvement of specific chemicals, pharmaceutical compounds and fragrances (Urlacher and Girhard, 2012). The p450 monooxygenase isolated from *Curvularia* sp, for instance, is currently used for biotransformation of Reichstein S to hydrocortisone (Sonomoto et al., 1983, Ghanem et al., 1992). Further, the oxidation of Compactin by a p450 monooxygenase from *Actinomadura macra* leads to the production of an effective drug (Pravastatin) for the reduction of blood cholesterol level (Ahmad et al., 2013). The production of erythromycin by *Saccharopolyspora erythraea* also requires a hydroxylation step performed by P450eryF (CYP107A1) (Park et al., 2010).

2.1.9.2 Bioremediation

Environmental contamination resulting from the release of crude oil and its products is one of the more common global environmental issues. The increasing demands for oil products in this century have led to an intensification of environmental pollution (Herwig, 2011). Crude oil consists of a mix of various hydrocarbons, including alkanes and aromatics, which are categorized as serious and hazardous environmental pollutants (Nolvak et al., 2012). The degree of toxicity of crude oil and its products depends on the chemical compounds present and their concentration (Nolvak et al., 2012, Anthony, 2006). However, many microorganisms are able to use these molecules as their source of carbon and thus are able to decrease their toxicity (Hao et al., 2004). Since oil contamination is a critical challenge for the environment, such studies are necessary for understanding the natural pathways involved in degradation of oil components and for the application of natural processes in oil bioremediation.

2.1.9.3 Biomarkers in oil biotechnology

Biomarkers are defined as those measurable activities or substances that represent a specific biological trait (Guermouche et al., 2013). In terms of ecological investigations, biomarkers are used for tracking any biological activity of interest such as the effects of toxic compounds or their assessment, source and distribution in an environment (Guermouche et al., 2013). Since biomarkers are able to measure the biological activity of interest before any tangible changes occur, abundant research has been performed focussing on the discovery of new biomarkers. For this purpose, several new and emergent technologies, such as mass spectrometry, metabolomics, metagenomics, genomics and proteomics have been applied to find new markers or a combination of agents that may be useful biomarkers. In terms of oil biotechnology, biomarkers are used for identification of petroleum deposits, determining the quality of oil, the efficiency of oil bioremediation and to identify the source of oil spills in the environment (Wang and Stout, 2010). Although oils are highly complex mixtures, varying from volatile molecules to very heavy aliphatic and aromatic hydrocarbons, their combination is unique from oil to oil (known as oil fingerprints) (Pauzi Zakaria et al., 2001). After the release of the mixture into an environment, its composition will change due to physical factors (dissolution, evaporation, degradation, photooxydation, oil–mineral aggregation (OMA), water–oil emulsification) and biological activities (biodegradation) (Oliveira et al., 2012, Shirani et al., 2012). A few days after oil spilling, the rate of loss through evaporation reaches 5-10% for heavy oils, 40% for petroleum products and 70% for light crudes (Lee et al., 2003). The biodegradation rate is mainly dependent on the nature of oil, the quantity of spilled oil, physical factors (such as temperature, pH and salinity) and the accessibility of nutrients, water and oxygen (Collier et al., 2012). An aged degraded oil sample is almost exhausted of branched and shorter straight chain n-alkanes and the concentration of PAHs is also much reduced (Yang et al., 2012). Some of the highly degradation-resistant molecules can be used for the analysis of the molecular changes that have occurred in the samples and for determining the fingerprints of sources (Yang et al., 2012). The selection of a suitable biomarker for oil analysis is dependent primarily on the nature of oil, conditions of the environment and purpose of the study (Shirani et al., 2012). In general, these compounds can be categorized into Volatile hydrocarbons (especially BTEX , benzene, toluene, ethyl methylbenzene, and Xylene), saturated aliphatic alkanes (C8 through C40), isoprenoids (such as pristane, phytane, trimethyl-C13, norpristane and farnesane), volatile paraffins and isoparaffins, cycloalkanes (mainly cyclopentane and cyclo-hexane compounds), PAHs and their alkylated (C1–C4) homologues, total petroleum hydrocarbons, total unresolved complex mixtures (UCM), stable carbon

isotope ratio ($d^{13}C$) and oxygen/nitrogen-containing heterocyclic hydrocarbons (Wang and Fingas, 2003, Wang and Stout, 2010). In addition to measurement of individual molecules, the ratios of some special compounds such as n-Cls/phytane and n-C17/pristane are used for tracking the effects of weathering on the mixture through times. Although, due to a similar volatility, physical factors do not change the ratios of pristane/phytane, n-Cas/phytane, and n-C17/pristane, biodegradation decrease the n-C1s/phytane, and n-C17/pristane ratios since microorganisms preferentially use short chain alkanes (Wang and Fingas, 2003). Furthermore, since PAHs are biologically converted into their corresponding alkylated homologues compounds, comparison of these compounds with their parents is used for determining the extent of oil weathering in a polluted site (Jacquot et al., 1996). Finally, since the amount of triterpanes and steranes are unique from oil to oil and because triaromatic steranes such as phenanthrene, dibenzothiophene and terpanes are more recalcitrant than other aromatic compounds they are used as effective biomarkers to investigate weathering events in oil and to track the origin of spilled oil (Wang et al., 1994, Xianqing et al., 2003).

2.1.10 Conclusion

Hydrocarbons are a group of chemical compounds consisting only of carbon and hydrogen atoms linked with covalent bonds that make them highly stable. Due to a lack of functional groups, these compounds show a weak solubility in water. Since the cyclic aromatic hydrocarbons are considered to be resistant to degradation due to their low reactivity, US EPA has classified these compounds as priority contaminants of natural resources. Because these molecules pose serious threat to the human and environmental health, understanding the catabolic pathways involved in the degradation of such molecules is of great research interest and also aids in designing efficient strategies for the bioremediation of hydrocarbon impacted environments. These persistent molecules affect niches of the habitat, thereby leading to the loss of biodiversity at both microscopic and macroscopic levels.

Although degradation pathways for many of both the aliphatic and aromatic hydrocarbons and several hydrocarbon degrading microorganisms from diverse phylogenetic groups have been identified, environmental contamination due to these compounds is still a big issue for environmentalists. Indeed, application of microorganisms or consortiums of different microorganisms with special abilities to degrade these compounds and even biotechnological approaches, have often not been successful in giving a significant reduction of these contaminants in a short time. This issue is worse when we have to clean promptly pollution in

soils and aquatic environments caused by oil tank spills. In addition to the type of habitat (aquatic or terrestrial) and climate, this failure can be ascribed to nutrient deficiency and the insolubility of hydrocarbons in water. Therefore, the microbial inoculum used for bioremediation of the polluted sites has to be supported by a strategy to resolve the problems. The inoculum can be supplied with additional nutrients, inorganic nitrogen and sulphur for instance, to support microbial growth. Furthermore the addition of a surfactant producing microorganisms or chemical biosurfactants can effectively resolve the unavailability of hydrocarbons due to their lack of solubility. The main problem in terms of unavailability is the presence of long chain hydrocarbons or their derivatives in the crude oil. Therefore, the microbial inoculum must contain those microorganisms with the ability to degrade these high molecular weight hydrocarbons. Although it is possible to construct genetically engineered organisms by combining the biosurfactant producing genes and all hydrocarbon degrading genes, including the genes responsible for degradation of short and long chain aliphatic hydrocarbons, branched hydrocarbons, cycloaliphatics and aromatics, the current rules for using genetic engineered microorganisms (GEM) prohibit commercial utilization of these microorganisms at least in near future. As a result, application of a consortium of microorganisms will be more effective and rational if each microorganism in this consortium is chosen based on its ability for the degradation of a special type of hydrocarbon or for facilitating this degrading activity. This review about the existing knowledge on degradation of different types of hydrocarbons by microorganisms is expected to help the researchers to further design and improve the technologies for bioremediation of the contaminants.

2.2 A review on the Genetics of Aliphatic and Aromatic Hydrocarbon Degradation

Published in; Applied biochemistry and biotechnology; 178 (2): 224-250.

2.2.1. Introduction

Due to the high diversity of the molecular structures of hydrocarbons, each class of compound has to be degraded by a specific enzyme (Abbasian et al., 2015a). Through the catabolic process, hydrocarbon molecules are first taken up by specific microbes, and then converted to simple organic molecules (Megharaj et al., 2011, Oliveira and De França, 2005). Based on the microbial species and the community in which the degrader species lives, these simple organic molecules may then be used by the same organism or may be released into the environment and further catabolized by other microbial members of the community (Gargouri et al., 2013). In terms of molecular investigations, such synergistic relationships suggest the probable presence of corresponding degradation genes in more than one member of the microbial community (Beškoski et al., 2011). This review describes the genes involved in the degradation of different types of aliphatic and aromatic hydrocarbons in both aerobic and anaerobic conditions.

All hydrocarbon oxygenases, including both monooxygenases and aromatic-ring-hydroxylating dioxygenases, are classified into a large family of iron-sulfur-flavoproteins containing enzymes capable of transferring one or two electrons to their substrates (van Beilen et al., 2006, Tinberg et al., 2011). Despite the existence of some monooxygenases that are structurally homomultimer (α_6) and physiologically related to each other (Herman et al., 2005), the aromatic ring hydroxylating dioxygenases usually have a heterohexamer ($\alpha_3\beta_3$) structure. Based on the sequence diversity of their α subunits, these aromatic hydroxylating dioxygenases are sub-classified into four groups: Toluene/Benzene oxygenase, Toluate/ Benzoate oxygenase, Naphthalene oxygenase and Biphenyl oxygenase (Schuler et al., 2009, Jouanneau et al., 2011). Furthermore, based on their native substrates the aromatic ring hydroxylating dioxygenases are classified into four groups: the toluene/biphenyl family (specialized for initial oxidative attack to benzene, toluene, chlorobenzenes, isopropylbenzene and biphenyl), the naphthalene family (for activation of naphthalene, phenanthrene, nitrotoluene and nitrobenzene), the benzoate family (for catabolism of toluate, benzoate, anthranilate,

isopropylbenzoate, trichlorophenoxyacetate, 2-chlorobenzoate) and finally the phthalate family (for the initial catabolism of aromatic acids such as phenoxybenzoate, *p*-toluene sulfonate, phthalate, vanillate, 3-chlorobenzoate) (Iwai et al., 2010, Iwai et al., 2011). In addition to these classified enzymes, several oxygenase enzymes have been identified which are specific for different substrates, such as salicylate, *o*-halobenzoate, 3-phenylpropionate, dibenzodioxin, aniline and dehydroabeitate (Gibson and Parales, 2000, Oberoi et al., 2015).

The full range of genes required for hydrocarbon degradation and their genetic organisation is not fully understood in many organisms. Although most of the genetic elements involved in the catabolism of aliphatic hydrocarbons are ordered in inducible operon structures with several co-regulated genes present in the same transcription unit, in some cases these operons are expressed constitutively (Cappelletti et al., 2011, Whyte et al., 2002a). Both inducible and constitutive genes can also be located on chromosomes either in the form of an operon unit or as separate genes (dos Santos et al., 2008, Rojo, 2009). Based on the catabolic genes present and their homology in endonuclease restriction patterns, DNA rearrangements and electrophoretic mobility, plasmids containing hydrocarbon degrading genes are divided into three groups: the plasmids containing alkane degradation genes (like the OCT plasmid) (Rojo, 2009), the plasmids containing naphthalene and salicylate degradation genes (such as the NAH plasmids) (Obayori and Salam, 2010) and finally the plasmids containing toluene and xylene oxidizing genes (like the TOL plasmids) (Obayori and Salam, 2010).

2.2.2 The alkane degradation genes

As an outline, biodegradation of hydrocarbons starts with oxidation of the substrate molecules by an electron carrier dependent monooxygenase system, producing the corresponding alcohol (Li et al., 2008b). After conversion of the hydrocarbons to their corresponding alcohol, the alcohols are further oxidized and broken down to smaller molecules that finally are utilized via the central catabolic reactions of the cells (reaction 2.3) (Rojo, 2010).

Based on the chain length of the aliphatic hydrocarbons utilised, n-Alkane utilising organisms are classified into three groups: methanotrophs, gaseous alkane (C2 to C4)-utilizing microorganisms and finally liquid alkane (C5 to C20)-catabolizing microorganisms (Tinberg and Lippard, 2011, Rojo, 2010). Furthermore, based on the molecular structure and the supporting electron transport system, monooxygenases are classified into rubredoxin dependent enzymes and (bacterial and fungal) cytochrome P450-containing monooxygenases

(Sayavedra-Soto et al., 2011, Girhard et al., 2010). The rubredoxin dependent enzymes are composed of a rubredoxin reductase, a rubredoxin and an alkane hydroxylase (Ji et al., 2013). In most bacteria, the rubredoxin dependent monooxygenases are encoded by the gene *alkB* while some bacteria, such as *Acinetobacter* sp., express the enzyme encoded by the *alkM* gene (Torres Pazmino et al., 2010).

Pseudomonas putida GP01, for instance, uses a monooxygenase to convert n-alkanes (C6-C10) into their fatty acids (Fig. 2.2.1) (Rojo, 2009). The *alk* gene organization in this strain is located on OCT plasmid, and encodes the enzymes required for degradation of C5-C13 *n*-alkanes (Rojo, 2009). In *P. putida* it is postulated that the *alkL gene* is involved in import of *n*-alkanes into the bacterial cells (Hearn E M et al., 2009). Furthermore, it has been shown that changes in the configuration of an outer membrane protein encoded by *blc* in *A. borkumensis* Sk2 can lead to the transport of short-length chain hydrocarbons into the cells (Lai et al., 2012). Furthermore, it is believed that the long-length chain fatty acid transporter proteins (*FadL*) in many bacteria participates in the transportation of long-length chain hydrocarbons into the cell (B, 2005). The alkanes are initially oxidized by a trimer alkane hydroxylase (a complex of alkane monooxygenase, rubredoxin and rubredoxin reductase encoded by *alkB, alkG* and *alkK*, respectively), which are integrated into the inner cell membrane of the bacterium via the product of *alkB* (Rojo, 2009). The resulting alcohol is further oxidized by the products of *alkJ* and *alkH* (respectively for an alcohol dehydrogenase and aldehyde dehydrogenase) into an aldehyde and acid, respectively, which is activated by addition of a CoA to the acid (through the action of the product of *alkK*). These genes are organised as *alkBFGHJKL* on the OCT plasmid and are controlled by the action of the products of another operon (*alkST*) located 40kb away from the first operon (Hernández-Arranz et al., 2013). The same operon structure (*alkBFGHJKL)* exists in *P. Putida* P1, but *alkST* has been moved to a position upstream of the operon and also the *alkL* and *alkN* genes are not separated by an insertion sequence (IS) (van Beilen et al., 2001). In *Acinetobacter* sp. strain ADP1. the three subunits of alkane hydroxylase (specialized for C6-C12 alkanes), alkane monooxygenase, rubredoxin and rubredoxin reductase, are encoded by *alkM, rubA* and *rubB*, respectively (Tanase et al., 2013). With the exception of *Rhodococcus eryithropolis,* none of the rubredoxin reductase genes are situated near an hydroxylase gene probably because of the involvement of the rubredoxin reductases in other metabolic pathways as well (Smits et al., 2002). This well organised aliphatic hydrocarbon degrading gene cluster is not always observed in all aliphatic hydrocarbon degrading bacteria. *Acinetobacter* HOI-N, for instance, is a hydrocarbon degrading bacterium

that contains a set of aliphatic hydrocarbon degrading enzymes located at three separate loci on the chromosome: the gene encoding the alkane hydroxylase is located at a considerable distance from the genes specified for the alcohol dehydrogenase as well as aldehyde dehydrogenase (Asperger and Kleher, 1991, Ratledge, 1984).

The alkane hydroxylating enzymes involved in initial activation of long-chain length aliphatic hydrocarbons (>C18) are evolutionary distinct from the previously mentioned enzymes involved in hydroxylation of short-length chain hydrocarbons. For instance, *AlmA*, encoded by *alma*, is a alkane monooxygenase belonging to the flavin-binding family in *Acinetobacter* sp. DSM17874, *Alcanivorax* and many other bacteria which involves in initial activation of the hydrocarbons with bigger than C32 (Wentzel et al., 2007). Furthermore, *LadA* is a thermophilic alkane hydroxylase, belonging to flavin-dependent oxygenase, obtained from *Geobacillus thermodenitrificans* NG80-2 with the ability to hydroxylate C15–C36 alkanes (Li et al., 2008a).

The genes involved in degradation of alkane hydrocarbons are down regulated by two regulatory systems, cytochrome ubiquinol oxidase (*Cyo*) and the global regulatory protein *Crc* (Dinamarca et al., 2003, (YusteandRojo,2001) to ensure expression of these genes just in certain physiological conditions. The *Cyo* gene product is known as a global regulatory factor able to regulate carbon metabolism and respiration. This factor suppresses the expression of alkane degrading genes in the presence of easily metabolised carbon sources (Dinamarca et al., 2003). The *Crc* gene product is a RNA-binding protein with the ability to stop the mRNA translation via binding to the 5′ end of the mRNAs responsible for production of both the regulatory factor *alkS* and alkane degrading proteins (Hernández-Arranz et al., 2013).

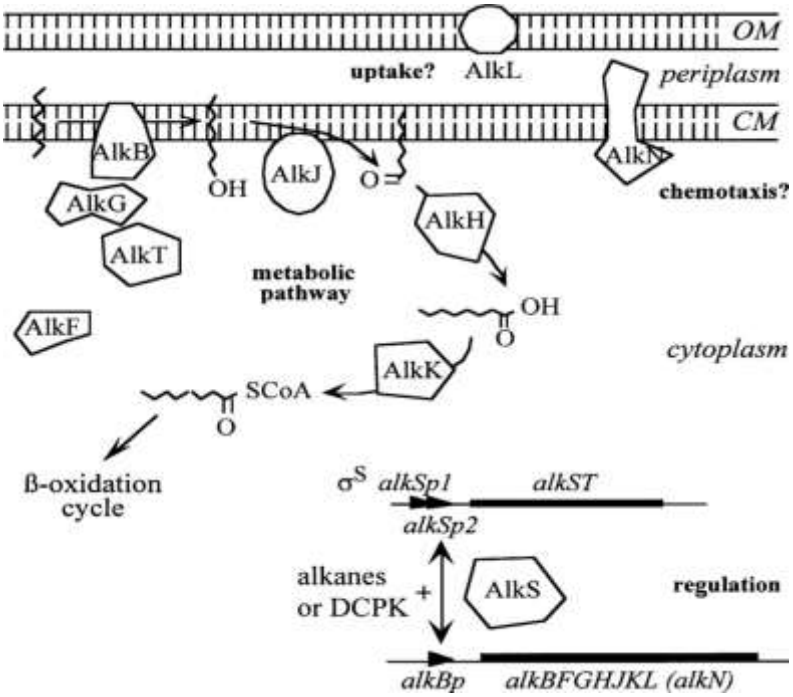


Figure 2.2.1. The position and role of alkane degrading proteins in *P. putida*. *alkB* = Alkane hydroxylase, *alkF*= Rubredoxin 1, *alkG*= Rubredoxin 2, *alkH*= Aldehyde dehydrogenase, *alkJ*= Alcohol dehydrogenase, *alkL*= Outer-membrane protein, *alkN*= Chemotactic transducer of *P. aeruginosa*, *alkT*= Rubredoxin reductase, *alkS*= Positive regulator (Van Beilen et al; 2001).

2.2.3 Cycloaliphatic compounds

The cycloaliphatic hydrocarbons, like cyclopentane, methylcyclopentane and cyclohexane are degraded by a large range of bacteria (Cheng et al., 2002). *Acinetobacter* strain SE19, for instance, uses six chromosomal catabolic genes, arranged as *chnBER* ORF and *chnADC* ORF for the degradation of cycloaliphatic hydrocarbons (Cheng et al., 2002). Through this reaction, cyclohexanone monooxygenase (encoded by *chnB*) and NAD(P)H-dependent aldehyde dehydrogenases (encoded by *chnE*) convert cyclohexane into cyclohexanone, which is further oxidized by the products of the *chnADC* ORFs. The *chnADC*, which is located in opposite direction of the *chnBER* ORF, encodes cyclohexanol dehydrogenase, 6-hydroxyhexanoic acid dehydrogenase and caprolactone hydrolase, respectively by *chnB*, *chnE* and *chnR*, to produce an end product of oxohexanoic acid. In the same way (Fig. 2.2.2), *R. ruber* SC1 uses a cyclododecane monooxygenase and a NAD(P)H-dependent aldehyde dehydrogenases to oxidaise cyclododecane first to cyclododecanol and then to cyclododecanone) (Kostichka et al., 2001a). A cyclododecanone monooxygenase, encoded by *cddA* oxidises it into a lactone

oxacyclotridecan-2-one (lauryl lactone), which is first hydroxlated by lauryl lactone esterase (encoded by *cddB)* to 12-hydroxydodecanoic acid and then is oxidized twice by two dehydrogenases (12-hydroxylauric acid dehydrogenase and 12-oxolauric acid dehydrogenase, encoded respectively by *cddC* and *cddD*, to make a 12-oxolauric acid and finally a DDDA (dodecanedioic acid) (Kostichka et al., 2001a). Although different bacteria utilize the same genes for the degradation of cyclododecane, their gene organisation may be different. In *R. ruber* SC1, the gene cluster is arranged as *cddABCDXY* with two spacesr ORFs between *cddABCD* and *cddXY*, while the gene order in *chn* cluster of *Acinetobacter* sp. strain SE19 is random (Kostichka et al., 2001a).

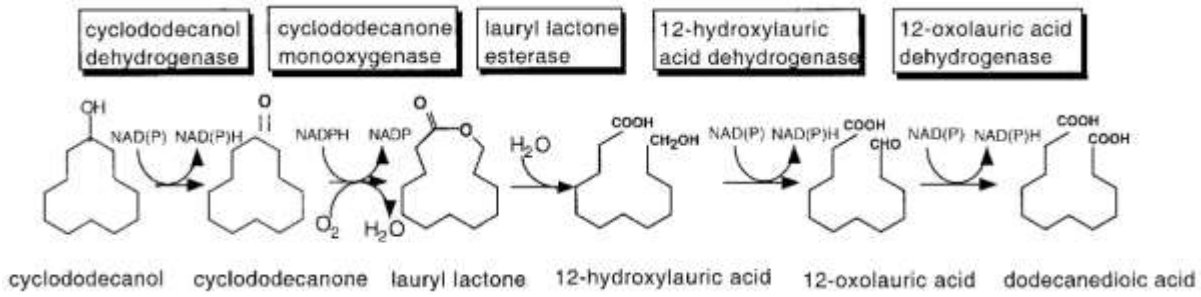


Figure 22.2. The metabolic pathway for degradation of cyclododecanol by *R. ruber* SC1 (Kostichka K. et al, 2001).

2.2.4 The plasmids containing naphthalene and salicylate degradation genes

Several aromatic degrading bacteria are able to convert mono/multiple cyclic aromatic hydrocarbons into salicylate, which undergoes a meta-cleavage to form the products to tricarboxylic acid cycle (TCA) (Izmalkova et al., 2013a). As a prototype dioxygenases enzyme, a (rieske-type two-iron two-sulphur centre containing) naphthalene dioxygenase (NOD; encoded by *nahAaAbAcAd*) inserts two oxygen atoms into the aromatic ring of a broad range of aromatic hydrocarbons, such as naphthalene, phenanthrene and anthracene, converting them to corresponding dihydrodiols, such as *cis*-naphthalene dihydrodiol and *cis*- phenanthrene dihydrodiol, respectively (Izmalkova et al., 2013b, Boronin and Kosheleva, 2010, Goyal and Zylstra, 1997). Next, a *cis*-dihydrodiol dehydrogenase (encoded by *nahB*) dehydrogenates the

dihydrodiols to make 1,2-dihydroxynaphthalene, which is subjected to *meta*-cleavage by 1,2-dihydroxynaphthalene dioxygenase (*nahC*) to form 2-hydroxychromene-2-carboxylic acid (Fig. 2.2.3). After an enzymatic *cis* to *trans* isomerization (by an isomerase encoded by *nahD*), the side-chain at the *trans* unsaturated bond of the *trans-o*-hydroxybenzylidenepyruvate product is cleaved by a hydratase-aldolase (encoded by *nahE*) to produce a salicylaldehyde. The product is finally dehydrogenated by NAD dependent salicylaldehyde dehydrogenase to salicylate (encoded by *nahF*). Depending on the bacterial strain, the covalent bond of the aromatic ring of salicylate is cleaved between two adjacent carbon atoms with hydroxyl groups (*meta*-cleavage) or between a carbon with a hydroxyl group and its adjacent carbon with a carboxyl group (ortho cleavage) (Seo et al., 2006). In most cases, like *P. putida* PpG7 (containing NAH7 plasmid) and *P. putida* R1 (containing SAL1 Plasmid), bacteria use a *meta*-cleavage reaction on salicylate (Fig. 2.2.4) (PhaLe et al., 2013, Yen et al., 1988) in which bacteria salicylate hydroxylase (*nahG*) convert salicylate into catechol. The product is oxidized by catechol oxygenase (*nahH*) to 2-hydroxymuconic semialdehyde. From here this intermediate can pass two different ways: in one way, the molecule is directly hydrolysed by a hydroxymuconic semialdehyde hydrolase (*nahN)* into 2-Oxo-4-pentenoic acid, while through a second pathway, the product is acted on by 2-hydroxymuconic semialdehyde dehydrogenase (*nahI*) and *4-oxalocrotonate isomerase (nahJ*) to produce 2-hydroxymuconic acid and 4-oxalocrotonic acid before conversion by 4-oxalocrotonate decarboxylase (*nahK)* into 2-Oxo-4-pentenoic acid. This intermediate is the substrate for 2-Oxo-4-pentenoate hydratase (*nahL)* and is converted to 4-hydroxy-2-oxovaleric acid which is broken by 2-Oxo-4-hydroxypentanoate aldolase (*nahM*) into pyruvic acid and acetaldehyde. The acetaldehyde is converted by Acetaldehyde dehydrogenase (*nahO*) into acetyl-CoA. In the ortho-cleavage pathway, on the other hand, bacteria use three subsequent enzymes, 2-oxo-4-hydtoxypentanoate aldolase *(nahM),* catechol 1,2-oxygenase (*carA*) *and* cis-muconate lactonizing enzyme (*catB*) to convert catechol into succinate and acetyl CoA (Fig. 2.2.4) (Van der Meer, 2008).

Through the gentisate pathway, bacteria employ an alternative pathway in which gentisate (2,5-dihydroxybenzoate) is subjected to a ring cleavage by gentisate1,2-dioxygenas (*BagI*) to produce a maleylpyrovate (Liu et al., 2011b, Agrawal et al., 1997). The product can directly hydrolyse into pyruvate and malate or may go into another process in which the product is first isomerised by an isomerase (*bagKL*) into fumarylpyruvate before hydrolysing (*bagK*) to pyruvate and fumarate (Liu and Zhou, 2012). Several other bacteria, such as *Salmonella*

typhimurium and *P. alkaligenese* are also able to use this system to degrade other substrates, such as 3-hydroxybenzoate and xylenol, respectively, through conversion to gentisate as an intermediate (Goetz and Harmuth, 1992, Feng et al., 1999).

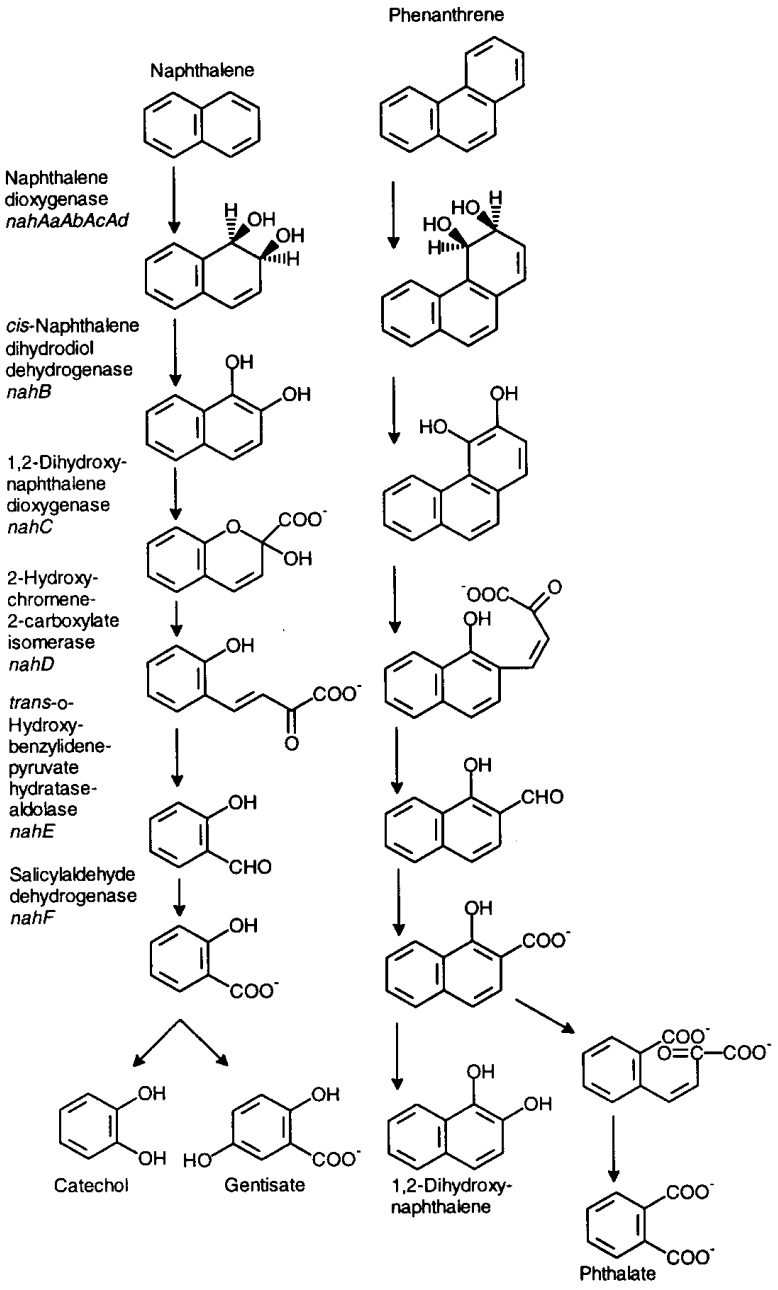


Figure 2.2.3. The metabolic pathway for degradation of naphthalene and phenanthrene (Goyal et al, 1997).

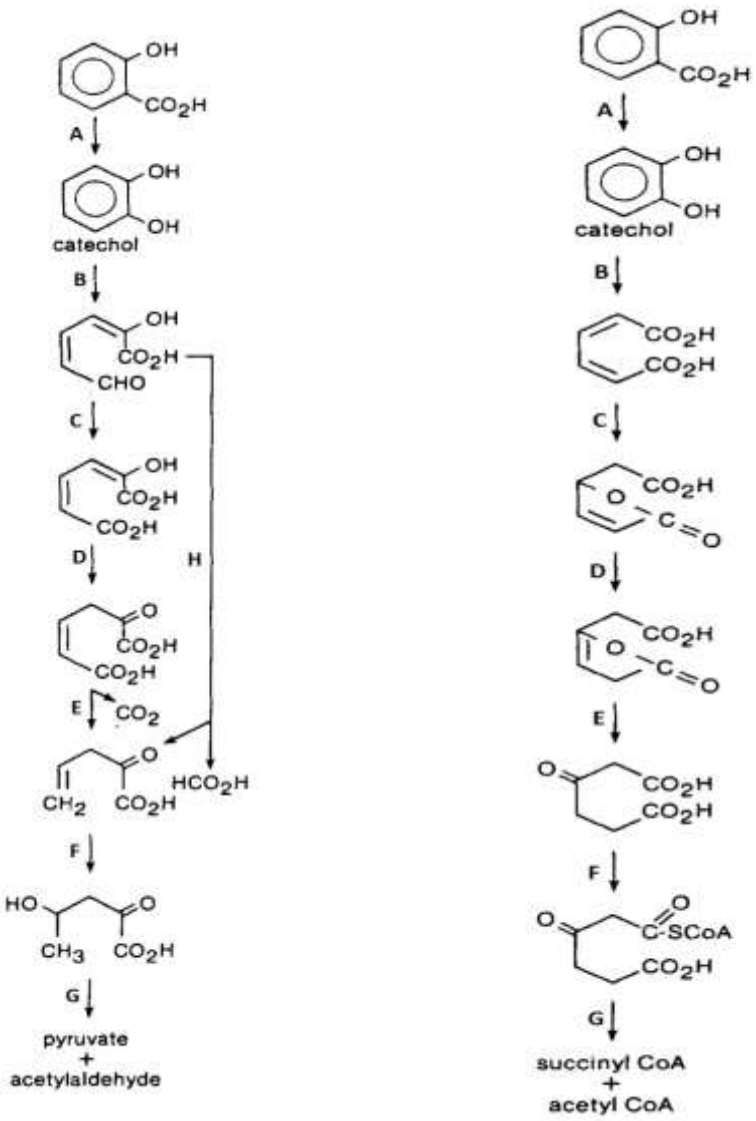
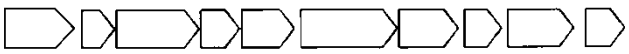


Figure 2.2.4. Degradation of catechol via meta-cleavage reaction (left) and orth-cleavage reaction (right) (From Yen KM and Serdar CM, 1998)

Although most of the genes responsible for degradation of naphthalene identified in different bacteria show 99–100% homology with their counterparts in other strains, the location (plasmid or chromosome) and organization of these gene clusters may be different in each strain (Fig. 2.2.5) (Li et al., 2004). However, there is a lower similarity between those genes identified in mycobacterial species and those in other bacteria probably due to the origin of the genes being from different sources or, less likely due to a greater rate of genetic changes in the

mycobacterial genes (Kim et al., 2008). NAH plasmids are a group of highly homologous plasmids, which carry naphthalene catabolic genes. These plasmids can be distinguished by their restriction endonuclease digestion patterns (Boronin and Kosheleva, 2014). The gene sequences of all identified NAH plasmids, such as NAH7 in *P. putida* PpG7 , pNL1 in *Novosphingobium aromaticivorans* F199, pND6-1 in *Pseudomonas* sp. strain ND6, and pWW60-1 in *P. putida* NCIB9816 are highly conserved, with 90-100% homology in the gene sequences (Obayori and Salam, 2010, PhaLe et al., 2013, Li et al., 2012). NAH7 in *P. putida* PpG1carries two separate operons of which the nah operon is specialized for conversion of PAHs, including naphthalene, anthracene and phenanthrene, to salicylate (*nahAaAbAcAdBFCED)* and the sal operon is used for the catabolism of salicylate to catechol and further to TCA cycle intermediates (*nahGTHINLOMKJ*) (Obayori and Salam, 2010). In addition, *nahX* (with an unknown function) and *nahY* (a chemotaxic transducer protein) are located downstream of the *nahJ* (Obayori and Salam, 2010, Pickrell et al., 2012). The product of the *nahY* gene acts as a methyl-accepting chemotaxis protein for cell attraction towards naphthalene via flagella-dependent movement (Pickrell et al., 2012). A *nahR* gene located between these two operons positively regulate the expression of both of the operons (Obayori and Salam, 2010, Habe and Omori, 2003).



Consensus	*nah*	*Aa*	*Ab*	*Ac*	*Ad*	*B*	*F*	*C*	*Q*	*E*	*D*
P. putida NCIB 9816-4	*nah*	*Aa*	*Ab*	*Ac*	*Ad*						
P. putida G7	*nah*	*Aa*	*Ab*	*Ac*	*Ad*			*C*	*Q*	*E*	*D*
P. putida BS202	*nah*	A1	A2	A3	A4	B	F	C		E	D
Pseudomonas sp. C18	*dox*		*A*	*B*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
P. aeruginosa PaK1	*pah*	*A*	*A2*	*A3*	*A4*	*B*	*F*	*C*	*Q*	*E*	*D*
P. putida OUS82	*pah*	*Aa*	*Ab*	*Ac*	*Ad*	*B*	*F*				
P. putida NCIB 9816	*ndo*		*A*	*B*	*C*						
P. fluorescens ATCC 17483	*ndo*		C1	C2	C3						
P. putida ATCC 17484	*ndo*		C1	C2	C3						

Figure 2.2.5. The gene organisation in different strains of *Pseudomonas* Sp. (Goyal, A and Zylstra G., 1997).

2.2.5 The plasmids containing toluene and xylene oxidizing genes

Through several pathways, bacteria insert one or more hydroxyl groups into aromatic rings to form catechol (Fig. 2.2.6), which is later cleaved for further catabolism. In one of these pathways, the toluene degrading genes in *P. putida*, located on the TOL plasmid, degrade this molecule into benzoic acid, cis-benzoate dihydrodiol and finally to catechol that in turn is cleaved for further oxidation processes (Fig. 2.2.6; Pathway A) (Zylstra and Gibson, 1991). Conversion of toluene into benzoate is performed by *xylA*, benzyl-alcohol dehydrogenase (*xylB*) and benzaldehyde dehydrogenase (*xylC*), while the next process, oxidation of toluate to catechol, is carried out by the products of the *xylD*, *xylE*, *xylF* and *xylG* genes. The first group of enzymes for the production of benzoate (encoded by *xylCAB*) is located on a plasmid, while the genes responsible for conversion of toluate into catechol can be found on both plasmids and chromosomes (Fig. 2.2.7) (Jutkina et al., 2013, Lehrbach et al., 1982, Shapiro et al., 1981). Catechol and its derivatives are cleaved via one of two *meta*-cleavage activities. In one pathway, 2-hydroxymuconic semialdehyde is directly converted via HMSH (hydroxymuconic semialdehyde hydrolase; encoded by *xylF*) into 2-oxopent-4-enoate or its derivatives (Harayama and Timmis, 2012). In the second pathway, 2-hydroxymuconic semialdehyde is first oxidized by HMSD (2-hydroxymuconic semialdehyde dehydrogenase (encoded by *xylG*) to its corresponding dioates before isomerization (encoded by *xylI*) to 2-oxopent-4-enoate (Fig. 2.2.8) (Harayama and Timmis, 2012). This last product is finally hydrolysed (by the *xyl*K encoded enzyme) to 4-hydroxy-2-oxovalerat before cleavage by 4-hydroxy-2-oxovalerate aldolase (encoded by *xylJ*) into pyruvate and propionaldehyde (Harayama and Timmis, 2012). The gene cluster on pTOL in *P. putida* is ordered as *xyl XYZLTEGFJQKIH* where *xylZ* and *xylL* encodes for 1,2-dioxygenase and 1,2-dihydroxycyclohexa-3,5-diene-carboxylate dehydrogenase, respectively and other genes downstream to *xylL* involve in the lower catabolic pathway (Jutkina et al., 2013).

Different strains of bacteria harbour quite similar operons in terms of DNA sequence and gene organization. *Sphingomonas yanoikuyae* B1 harbours a *xylXYEFGJQKIHT* operon where in addition to internal gene rearrangements, the *xylL and xylZ* genes have moved to a separate place on the genome (Kim and Zylstra, 1999). Furthermore, the operon *bphR1-bphA1A2(orf3)bphA3A4BCX0X1X2X3D* in *Burkholderia* sp. strain LB400 and *P. pseudoalcaligenes* KF707, responsible for degradation of biphenyls to pyruvate and acetaldehyde, is highly homologous to the operons for degradation of toluene (Fig. 2.2.9)

(Furukawa et al., 2004). The corresponding operon in *Pseudomonas* KKS102 is broken into a *bphEGF*(ORF4)*A1A2A3BCD*(ORF1)*A4* operon, which allows the catabolic reaction of biphenyls to proceed to 2-hydroxypenta-2,4-dienoate and benzoic acid, and the gene cluster *bphEGF* located 4kbp upstream of the first operon, which encode for hydratase, aldolase and dehydrogenase and convert these intermediate products into pyruvate and acetaldehyde (Ohtsubo et al., 2000). Finally, the gene cluster for the degradation of biphenyl in *Rhodococcus* sp. RHA1 is distributed between several linear plasmids, referred to as RHA1, 2 and 3. Most of the genes for initiating the catabolism of biphenyl (*bphA1A2A3A4CB*) are located on RHA1, while the *bhpDEF* cluster is placed on pRHL2 (Takeda et al., 2004).

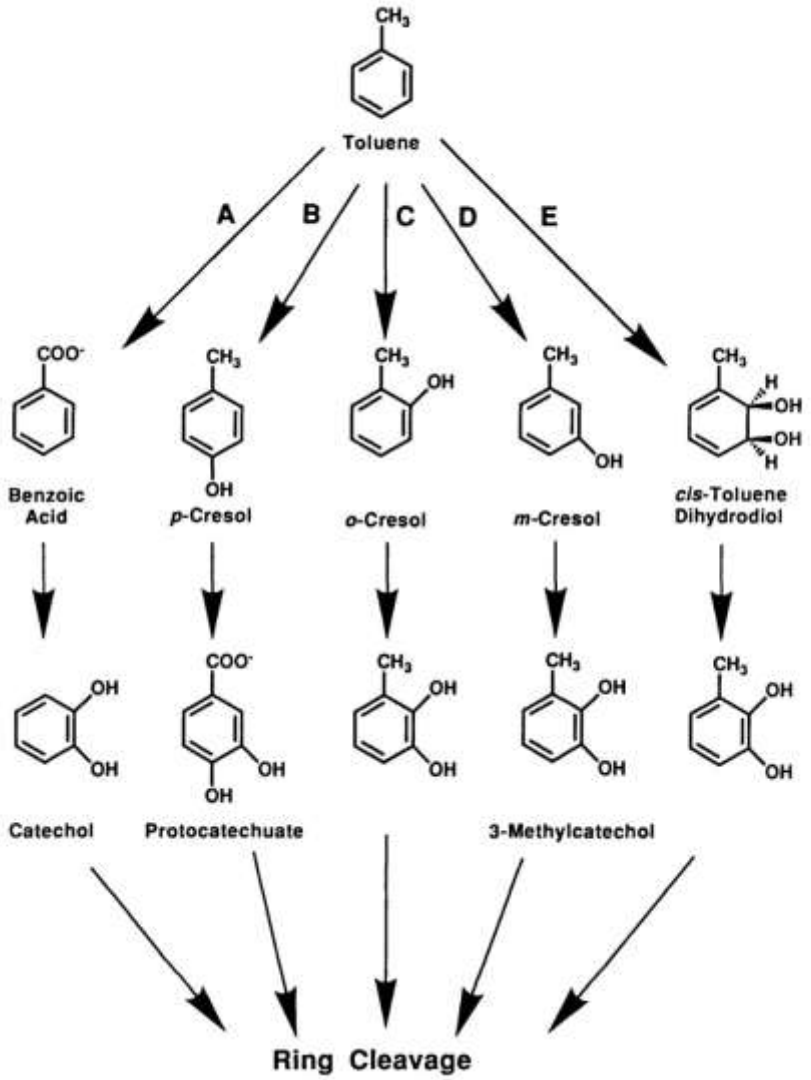


Figure 2.2.6. Five possible pathways for degradation of toluene (Zylstra GJ and Gibson DT, 1991).

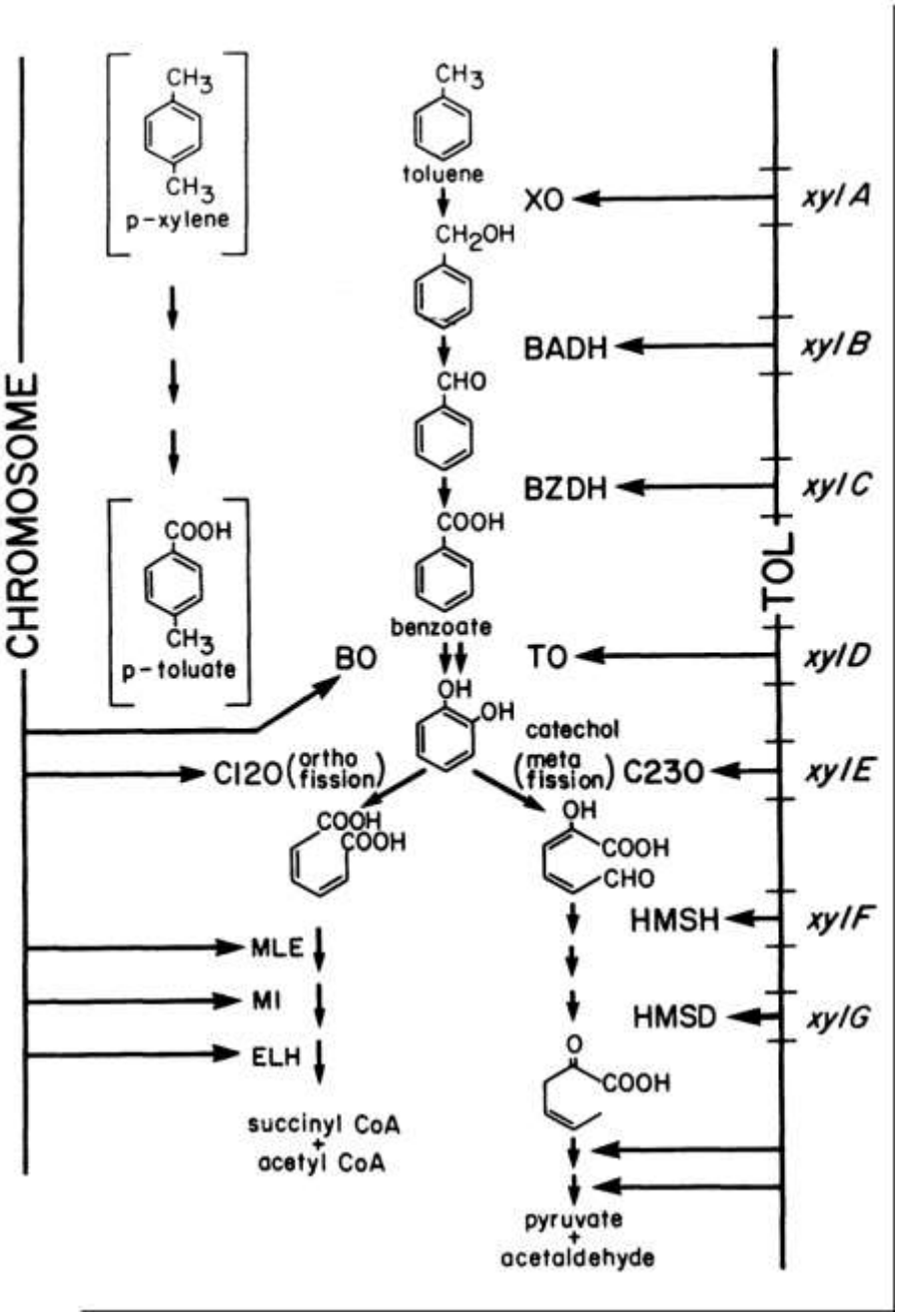


Figure 2.2.7. Gene organisation of toluene and xylene degradation on chromosome and TOL plasmid of *P. putida*. XO = xylene oxidase; BADH = benzyl alcohol dehydrogenase; BZDH = benzaldehyde dehydrogenase (Shapiro J. et al, 1981)

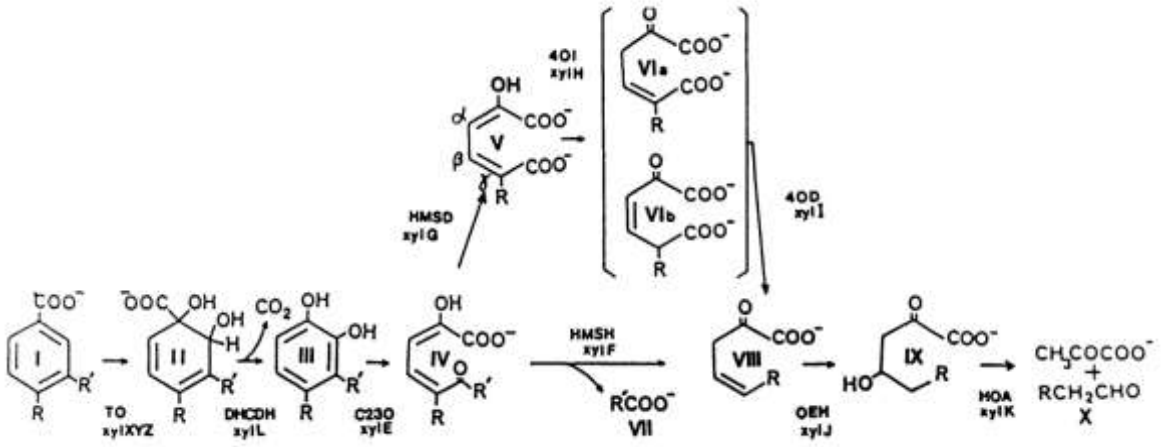


Figure 2.2.8. The metabolic pathway for the degradation of toluate and benzoate via the meta-cleavage process. The 2-hydroxymuconic semialdehyde intermediate can oxidise via two different reactions. It may be oxidised directly by HMSH ((hydroxymuconic semialdehyde hydrolase) into 2-oxopent-4-enoate or may be first converted by HMSD (2-hydroxymuconic semialdehyde dehydrogenase into a dioates before isomerization to 2-oxopent-4-enoate. TO, toluate 1,2-dioxygenase; DHCDH, 1,2-dihydroxycyclohexa-3,5-diene-carboxylate dehydrogenase; C230, catechol 2,3-dioxygenase; 40D, 4-oxalocrotonate decarboxylase; OEH, 2-oxopent-4-enoate hydratase; HOA, 4-hydroxy-2-oxovalerate aldolase. xylE to xylZ are names of the genes specified for the degrading enzymes (Harayama S., 1987).

Through a completely different pathway in *P. putida* Fl (Figures 2.2.6D and 2.2.10) a multimeric enzyme referred to as toluene dioxygenase (encoded by tod *C1C2BA*) converts toluene and many other aromatics into (+*)-cis*-1(S),2(R)-dihydroxy-3-methylcyclohexa-3,5-diene (Choi et al., 2003, Shindo et al., 2005). This reaction is driven by the protons and electrons originating from NADH that are passed through an electron transport system composed of a reductase (encoded by *todA*), a ferrodoxin (encoded by *todB*) and an Iron Sulphur Protein or ISP (encoded by *todC1* and *todC2*). A NAD^+-dependent-Cis-toluene dihydrodiol dehydrogenase (encoded by *todD*) oxidises the dihydrodiol to form 3-methylcatechol, which is cleaved twice by 3-methylcatechol-2,3-dioxygenase (encoded by *todE*) to 2-hydroxy-6-oxo-2,4-heptadienoate and then by 2-Hydroxy-6-Oxo-2,4-hepladlenoale

hydrolase (encoded by *todF*) to 2-hydroxypenta-2,4-dienoate and acetate. All of the toluene degrading genes are ordered as *todFClC2BADE* gene cluster.

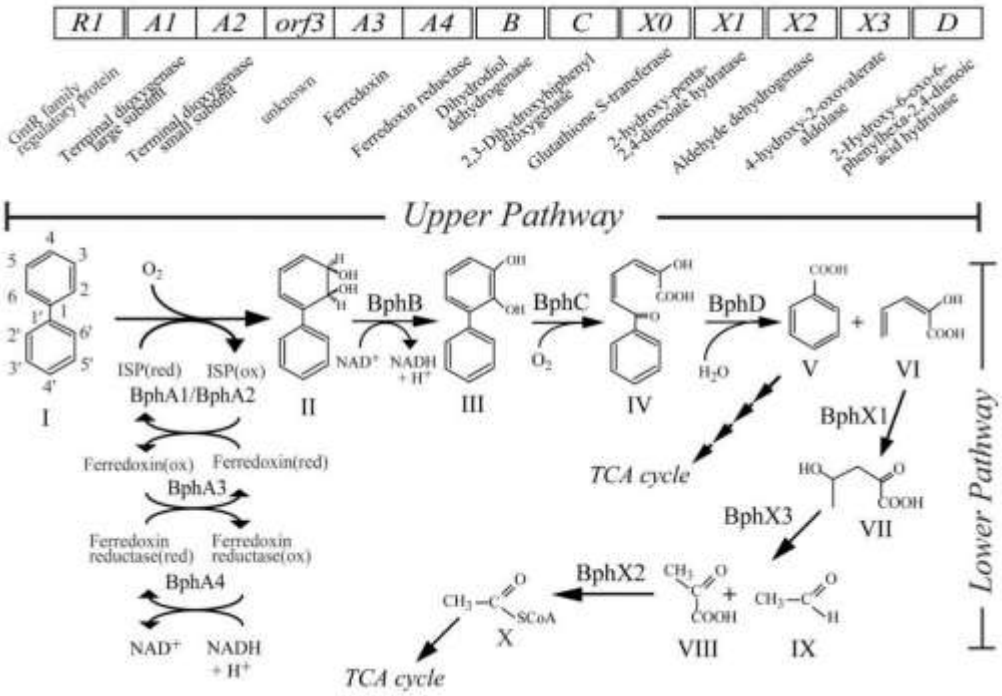


Figure 2.2.9. The degradation pathway of biphenyl into acetyl CoA and the order of the corresponding genes (*bph* operon) in *P. pseudoalcaligenes* KF707. The biphenyl dioxygenase (consisting of two subunits of the terminal dioxygenase and ferrodoxin encoded respectively by bphAlA2A3A4) inserts two oxygen atoms into biphenyls (I) to yield dihydrodiols (II). The product is further oxidised by dihydrodiol dehydrogenase (encoded by bphB) to 2,3-Dihydroxybiphenyl (III), which undergoes a cleavage at the meta-position by the product of the bphC gene (2,3-dihydroxybiphenyl dioxygenase) to yield 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate (IV). This molecule is broken by 2-hydroxy-6-oxo-6phenylhexa-2,4-dienoic acid hydrolase (encoded by bphD) into benzoic acid (V) and 2-hydroxypenta-2,4-dienoate (VI). A hydratase, encoded by (bphX1) converts the 2-hydroxypenta-2,4-dienoate to 4-hydroxy-2-oxovalerate (VII), which is cleaved by 4-hydroxy-2-oxovalerate aldolase (encoded by bphX3) into pyruvate (VIII) and acetaldehyde (XI). The acetaldehyde is then activated by acetaldehyde dehydrogenase (encoded by bphX2) and glutathione-S-transferase (encoded by bphX0) through the addition of Co-A (X) before entering the TCA cycle. ISP, iron-sulfur protein (Furukawa K, 2004).

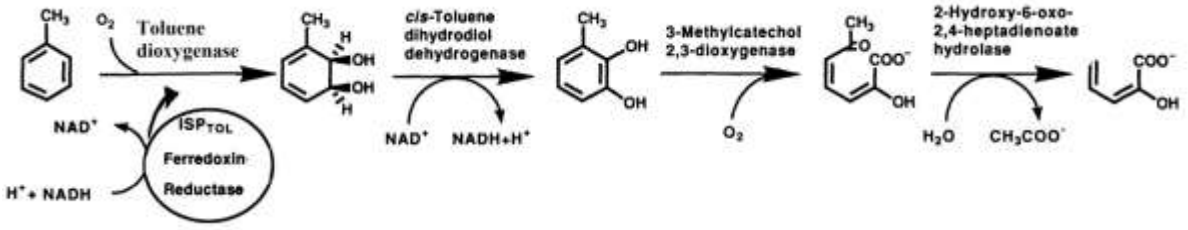


Figure 2.2.10. Conversion of toluene into 2-hydroxypenta-2,4-dienoate by *P. putida* F1. The electron originating from NADH, and H^+ is passed through an electron chain system, respectively consisting of reductase, ferrodoxin and an Iron sulphur protein (ISP) to catabolise toluene into toluene dihydrodiol dehydrogenase (Shindo K, 2005).

2.2.6 Nitroaromatic compounds

Nitroaromatic compounds (NACs) are synthetic molecules broadly utilized in different industries as plastics, pharmaceuticals, precursors for dyes, explosives and pesticides (Halecky et al., 2013). Although there are many different types of nitroaromatics, 2,4,6-trinitrotoluene (TNT), dinitrotoluenes and nitrotoluenes are the most abundant environmental pollutants (Bińet al., 2001, Funk et al., 1995). Based on the gene capacity and type of the original nitroaromatic compound, microorganisms use oxidative and/or reductive degrading pathways to convert these NACs completely to CO_2 and H_2O or partially to an organic compound (Arora et al., 2014). While aerobic bacteria use both the catabolic systems, anaerobic bacteria are able to use only the reductive degrading mechanism to catabolize NACs (Arora et al., 2014).

The oxidative reactions are triggered through the reaction of a mono/di-oxigenase enzyme, releasing a nitrite and dihydroxy aromatic compounds. The substrate specificity and the intermediate and final products are unique based on the substrate, the type of oxygenases used in the reaction and the organisms involved in the degradation. The monooxygenase systems, for instance, are able to react with different substrates, including 2-nitrophenol (*P. putida* B2**)** (Zeyer et al., 1986), 4-nitrophenol (*Moraxella* sp.) (Spain and Gibson, 1991) and 4-methyl-5-nitrocatechol (*Pseudomonas* sp. strain DNT) (Haigler et al., 1994), 2-nirrotoluene (from *Acidovorax* sp. JS42) (Lee et al., 2005), nitrobenzene (*Comamonas* sp. strain JS765) (Nishino and Spain, 1995), 3-nitrobenzoate (51), 1,3-dinitrobenzene (*Burkholderia cepacia* R34)

(Johnson G. R. et al., 2002), 2-chloronitrobenzene (*P. stutzeri* strain ZWLR2-1) (Liu et al., 2011a) and 2,4-dinicrotoluene (Spanggord et al., 1991). The monooxygenases belonging to the TC-FDM (a two-component flavin-diffusible monooxygenase) family in *Moraxella* sp., *Pseudomonas* sp. strain ENV2030, *Rhodococcus* sp. strain PN1, *Rhodococcus opacus* SAO101 and many other bacteria, oxidizes 4-nitrophenol in expense of two NADPH and a molecular oxygen to hydroquinone and releases a nitrite molecule (Perry and Zylstra, 2007). The members of this family can be divided into two homology groups: the phenol 2-monooxygenase and phenol 4-monooxygenase groups. While the members of first group (such as phenol monooxygenase (*PheA*), nitrophenol monooxygenase (*NphA1*) and 4-hydroxyphenylacetate monooxygenase (*HpaB*), hydroxylate the ortho group of phenols, members of the second group, including 2,4,6-trichlorophenol monooxygenases (*TcpA*), 2,4,5-trichlorophenol monooxygenase (*TftD*), PNP monooxygenase (*NpcA*) and 4-chlorophenol monooxygenase (*CphC*-I), hydroxylate their para position. Following release of nitrite in both bacteria, the products are directed into normal cell metabolism that lead to production of maleylacetic acid and further to β-ketoadipate. The npd gene cluster in *Arthrobacter* sp. JS443, responsible for catabolism of p-nitrophenol consists of three genes, *npdB* (hydroxyquinol 1,2-dioxygenase), *npdA1* (p-nitrophenol monooxygenase), and *npdA2* (p-nitrophenol hydroxylase) (Perry and Zylstra, 2007). This cluster is 85% similar to the cph I gene cluster found in *A. chlorophenolicus* A6 (Fig. 2.2.11) (Perry and Zylstra, 2007). However, while *cph* gene cluster are regulated by products of two genes (cphR and cphS), these regulatory genes are combined into a single gene, called as *npdR* (Nordin et al., 2005, Perry and Zylstra, 2007).

The multicomponent NACs dioxygenase system, , such as NBDO, NDO, DNTDO, 3-NTDO, 2NTDO and TNT dioxygenase, consists of a terminal iron-sulfur oxygenase, an iron-sulfur ferredoxin and a flavoprotein reductase to substitute a hydroxyl group to the ring by a nitrite (Singh et al., 2014b). NACs dioxygenases are usually nonspecific enzymes able to react with several NACs, such as nitrotoluenes, dinitrobenzenes and nitrophenols and non-nitrogen aromatic hydrocarbons as well. These dioxygenases use the electrons transferred by two other components to the system to add two oxygen atoms into the ring to produce catechol intermediates (Singh et al., 2014b). The dioxygenase genes show highly similarity with the sequence and gene structure of the naphthalene, which belongs to Rieske non heme iron dioxygenases (Ferraro et al., 2005). The genes responsible for expression of these subunits are organised in a sequence of a reductase (*mntAa*), followed by one or two other ORF, the ferredoxin subunits, the large subunit of oxygenase (*MntAc*)and its small subunit (*MntAd*).

These genes are under control of a regulator (*mntR*) located at upstream of the functional dioxygenase gene. Although the two ORFs between genes reductase and ferredoxin in many strains are inactive, in several cases, such as DNTDO identified in *Burkholderia* sp. strain DNT and NDO identified in *Ralstonia* sp., are responsible for expression of two subunits of salicylate-5-hydroxylase, which accept the electron released from reductase and ferredoxin. As a prototype NACs, 2,4-Dinitrotoluene (2,4-DNT) (Figures 2.2.12 and 2.2.13) is initially oxidized by a tetramer dioxygenase (encoded by *dntAaAbAcAd*) to release NO_2^-, converting it to 4-Methyl-5-nitrocatechol (MNC). While in *Acidovorax* sp. strain JS42 this operon is regulated by a transcriptional activator (*NtdR*) located immediately upstream of the operon (Rabinovitch-Deere and Parales, 2012), there is a gap between the *dntAaAbAcAd* and its upstream regulator (*dntR*) in *Burkholderia cepacia* R34 (Symons and Bruce, 2006). This intermediate is affected by the MNC monooxygenase (encoded by *dntB*) to produce 2-Hydroxy-5-methylquinone (HMQ), which is further oxidised by HMQ reductase (encoded by *dntC*) to 2,4,5-Trihydroxytoluene (THT). THT is undergone a extradiol ring fission by THT oxygenase (encoded by *dntD*) to produce 2,4-Dihydroxy-5-methyl-6-oxo-2,4-hexadiennoic acid (DMOH). This intermediate is converted by bifunctional DMOH isomerase (encoded by *dntG*) first to 4-Hydroxy-2-keto-5-methyl-6-oxo-3-hexenoic acid and further to pyruvic acid and Methylmalonic acid semialdehyde. The last product is bound by a CoA-dependent methylmalonate semialdehyde dehydrogenase (encoded by *dntE*) to CoA-SH in conjunction with NAD^+ to produce propionyl-CoA.

2.2.7 Dibenzothiophene (DBT)

Since a carbon group of Dibenzothiophene (DBT) and its derivatives is substituted with sulfur, their degradation requires an initial set of reactions to remove this sulfur (Lu et al., 2007). The genes involved in desulfurization of Dibenzothiophene (*DszA*, *DszB* and *DszC*) are located on an operon. These genes remove the sulfur through a four step reaction referred to as the 4S desulfurization pathway (Raheb, 2011, Shavandi et al., 2013) (Fig. 2.2.14). The two first reactions are catalysed by the action of Dibenzothiophene monooxdase (*DszC*) in which DBT is converted initially to Dibenzothiophene-sulfoxide and later to Dibenzothiophene sulfone ($DBTO_2$) (Abin-Fuentes et al., 2013, Martinez et al., 2015). The product of *DszA* ($DBTO_2$ monooxygenase) converts this intermediate to 2-(2-hydroxyphenyl)-benzene sulfonate (HBPS) (Abin-Fuentes et al., 2013, Martinez et al., 2015), which then is more metabolised by HBPS-desulfinase (*DszB*) releases sulfur from the compound and leaves 2-Hydroxybiphenyl (HBP). The first three steps are O_2-dependent oxidative reactions and require an electron and hydrogen

transportation system, composed of $FMNH_2$ and NAD(P)H as a reductant (Abin-Fuentes et al., 2013, Martinez et al., 2015). Oxidation of NAD(P)H to NAD(P) is catalysed by action of a flavin reductase, which is encoded by *DszD* (Abin-Fuentes et al., 2013, Martinez et al., 2015). This later genes is located at a separate locus from *dszABC* (Abin-Fuentes et al., 2013, Martinez et al., 2015). Following desulfurization of these compounds, the product is a simple aromatic compound which is metabolised by the aromatic degradation system as mentioned before. Although there is not enough evidence for the regulation of these genes, it has been shown that expression of the operon is under the control of a repressor and is limited in the presence of different readily bioavailable sulfur sources, such as SO_4^{2-}, Casamino acids, methanesulfonic, taurine, cysteine and methionine (Mohebali Gasemali and Andrew, 2008).

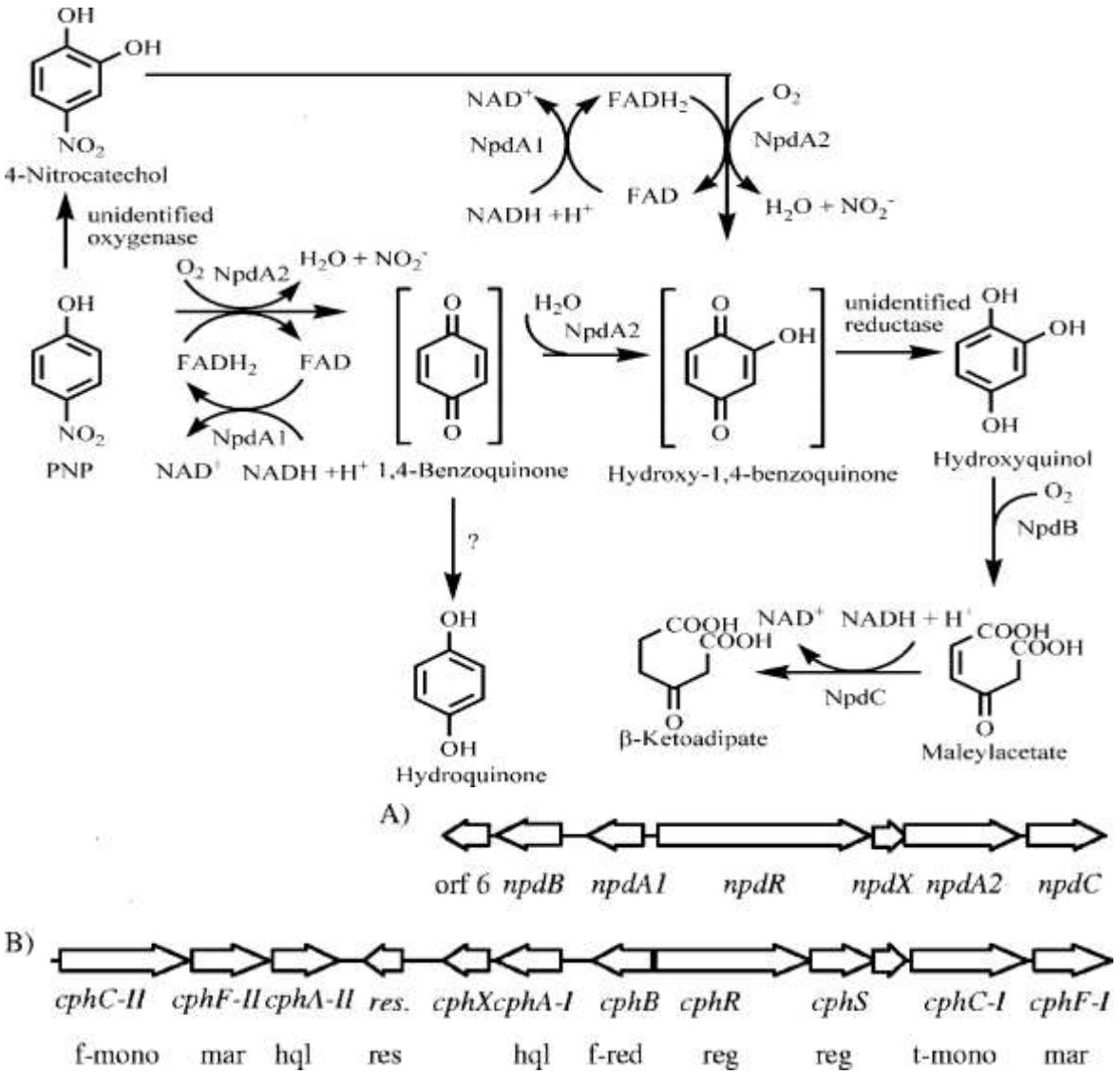


Figure 2.2.11. The metabolic pathway of p-nitrophenol (PNP) and the npd gene cluster in *Arthrobacter* sp. JS443 (A) and *A. chlorophenolicus* A6 (B); f-mono, flavoprotein monooxygenase; mar, maleylacetate reductase; hql, hydroxyquinol 1,2-dioxygenase; res, resolvase pseudogene; f-red, flavin reductase; reg, transcriptional regulator; t-mono, TC-FDM. (Perry LL and Zylstra GJ, 2007).

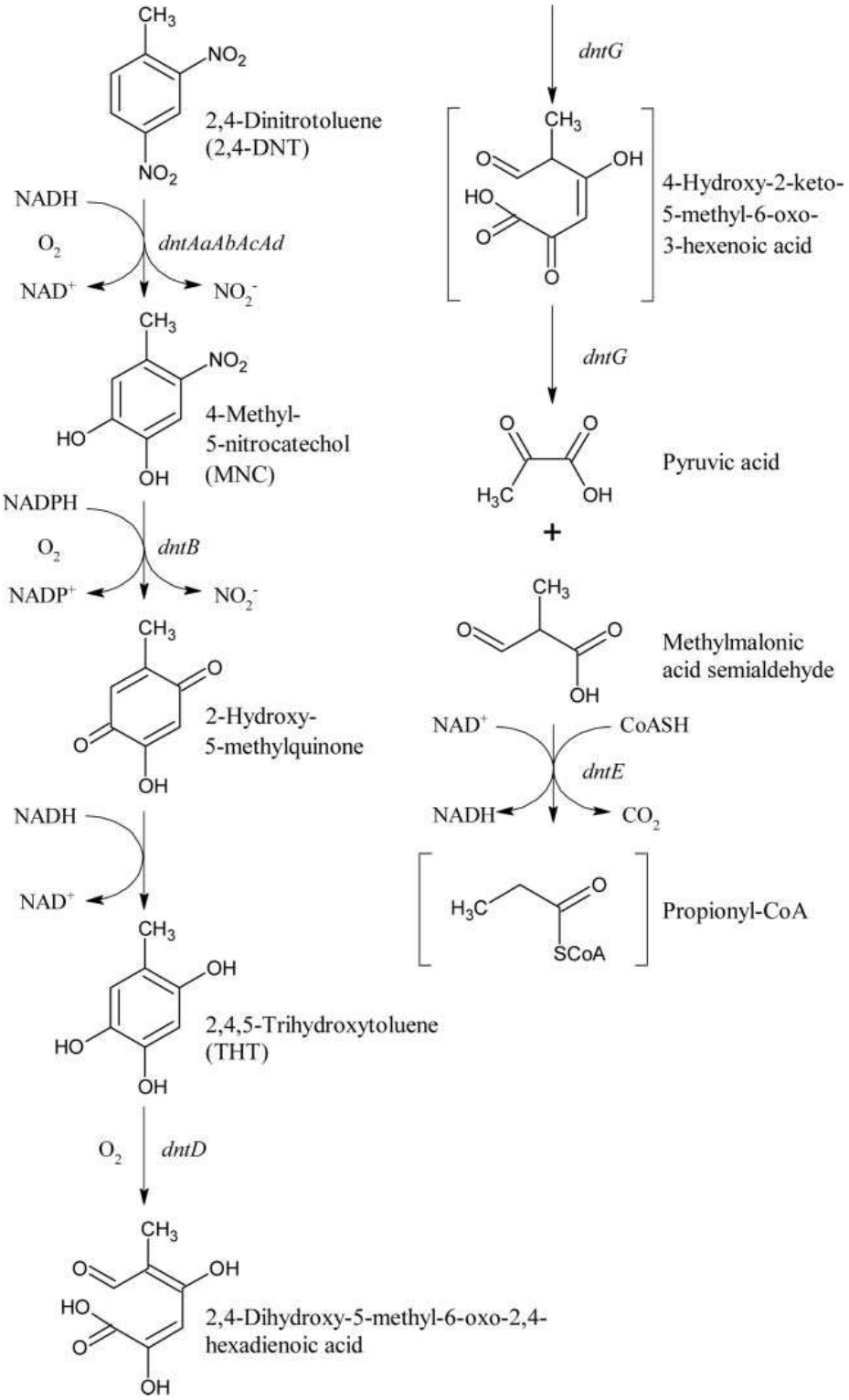


Figure 2.2.12. Oxidative catabolism of 2,4-dinitrotoluene in *Burkholderia cepacia* R34. *dntAaAbAcAd* (2,4-DNT dioxygenases); *dntB*: (methylnitrocatechol monooxygenase); *dntD* (the extradiol ring fission enzyme); *dntG* (bifunctional isomerase/hydrolase); *dntE* (CoA-dependent methylmalonate semialdehyde dehydrogenase) (Symons, Z.C. and N.C. Bruce, 2006).

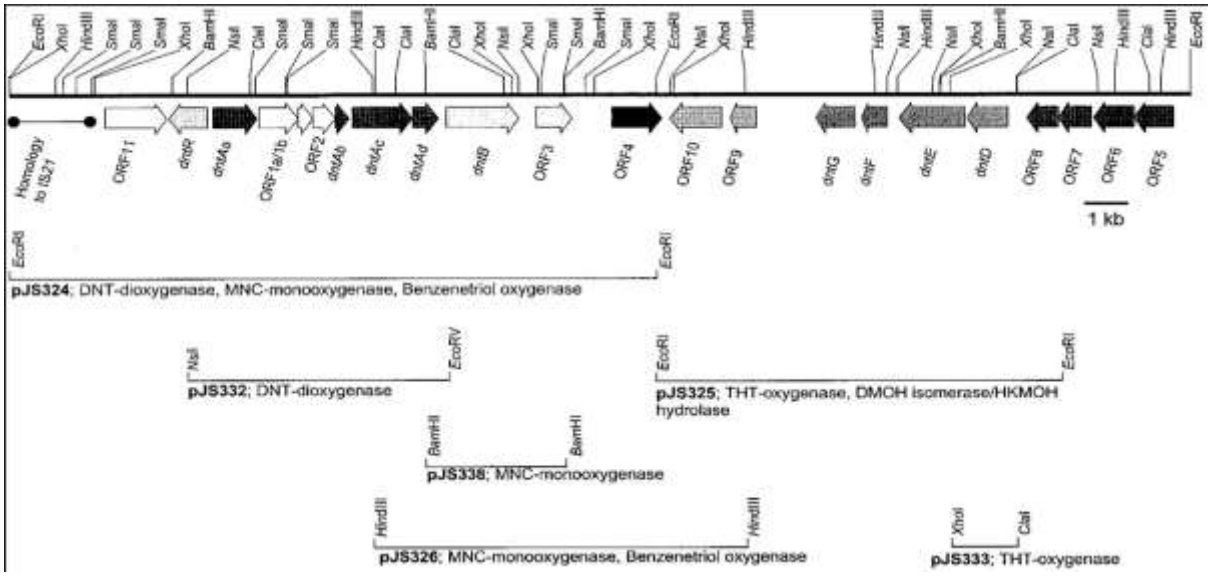


Figure 2.2.13. Physical map of region encoding 2,4-DNT pathway genes in *Burkholderia cepacia* R34 (Jahnson et al., 2002).

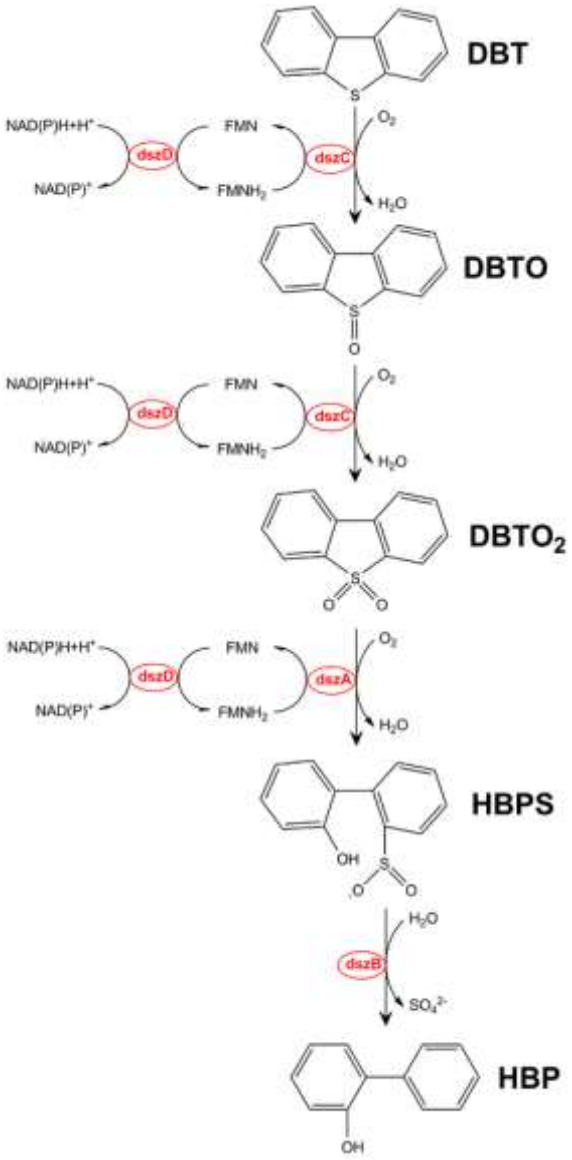


Figure 2.2.14. The Dibenzothiophene desulfurizing pathway in *Gordonia* sp; *dszA*: Dibenzothiophene sulfone ($DBTO_2$) monooxygenase, *dszB*: 2-(2-hydroxyphenyl)-benzene sulfonate (HBPS) desulfinase, *dszC*: Dibenzothiophene monooxdase and *dszD*: NAD(P)H:FMN oxidoreductase, DBT: Dibenzothiophene, DBTO: Dibenzothiophene-sulfoxide, $DBTO_2$: Dibenzothiophene sulfone, HBPS: 2-(2-hydroxyphenyl)-benzene sulfonate, HBP: 2-Hydroxybiphenyl

(taken from http://2012.igem.org/File:12SJTU_desulpathway1.png).

2.2.8 Anaerobic degradation of hydrocarbons

Several phylogenetically and physiologically distinct microorganisms degrade hydrocarbons through anaerobic metabolic pathways utilising the reduction of unusual electron acceptors, such as sulphate, thiosulfate, nitrate, nitrite, nitrous oxide, metal ions and carbonate), or using anoxygenic phototrophic reactions involving the donation of electrons and hydrogen for substrate catabolism activities (Grossi et al., 2008). These anaerobic bacteria degrade hydrocarbons via five different pathways: a) Addition of a fumarate to methylene or methyl groups of hydrocarbons (von Netzer et al., 2013), b) Oxygen-independent hydroxylation of 2th or 3th terminal C-atoms (to make secondary or tertiary alcohols) (Szaleniec et al., 2007), c) Carboxylation of un-substituted carbon atoms of aromatics (Broberg and Clark, 2010), d) Hydration of the double and triple bond of alkenes and alkynes (Liu et al., 2011c) and e) Reverse methanogenesis (Thauer, 2011).

Many anaerobic bacteria, including denitrifying microorganisms, sulphate reducing bacteria, methanogenic consortia and metal (Mn IV, Fe III) reducing bacteria, are able to activate hydrocarbons via terminal or sub-terminal addition of a carbonic group, such as fumarate, to a carbon atom of the hydrocarbon (Selesi and Meckenstock, 2009b). Conversion of toluene into (R)-benzylsuccinate is a common example in which a trimer benzylsuccinate synthase (BSS) enzyme (encoded by *bbsABC*) adds a fumarate to the substrate (Fig. 2.2.15) (von Netzer et al., 2013, Leuthner and Heider, 2000b). Then, after addition of a CoA to the product via the action of succinyl-CoA:benzylsuccinate CoA-transferase (encoded by *bbsEF*), the benzylsuccinyl-CoA undergoes an oxidation step, a hydration and another oxidation respectively by benzylsuccinyl-CoA dehydrogenase (encoded by *bssG*), phenylitaconyl-CoA hydratase (encoded by *BbsH*) and (5)3-hydroxyacyl-CoA dehydrogenase (encoded by *BbsCD*) before cleavage (by benzoylsuccinyl-CoA thiolase (encoded by *BbsB*) into a benzyl-CoA and a succinate. All of these genes are clustered as a *bbsABCDEFGHI* operon (*bbs* is abbreviation of beta-*oxidation* of benzylsuccinate) and is controlled by a regulatory factor (TdiSR) that activates this operon in the absence of O_2 and the presence of toluene (Hermuth et al., 2002). The Naphthyl-2-methyl-succinate synthase in SRBs, encoded by *nmsABC*, specific for transfer a fumarate to naphthalene is similar to the corresponding enzyme involved in the biodegradation of toluene (Selesi et al., 2010). Furthermore, the smallest subunit of alkylsuccinate synthases, involved in addition of a fumarate to alkanes through anaerobic

degradation, is highly similar to the corresponding subunit of benzylsuccinate synthetase, encoded by *bssC* (Ehrenreich et al., 2000).

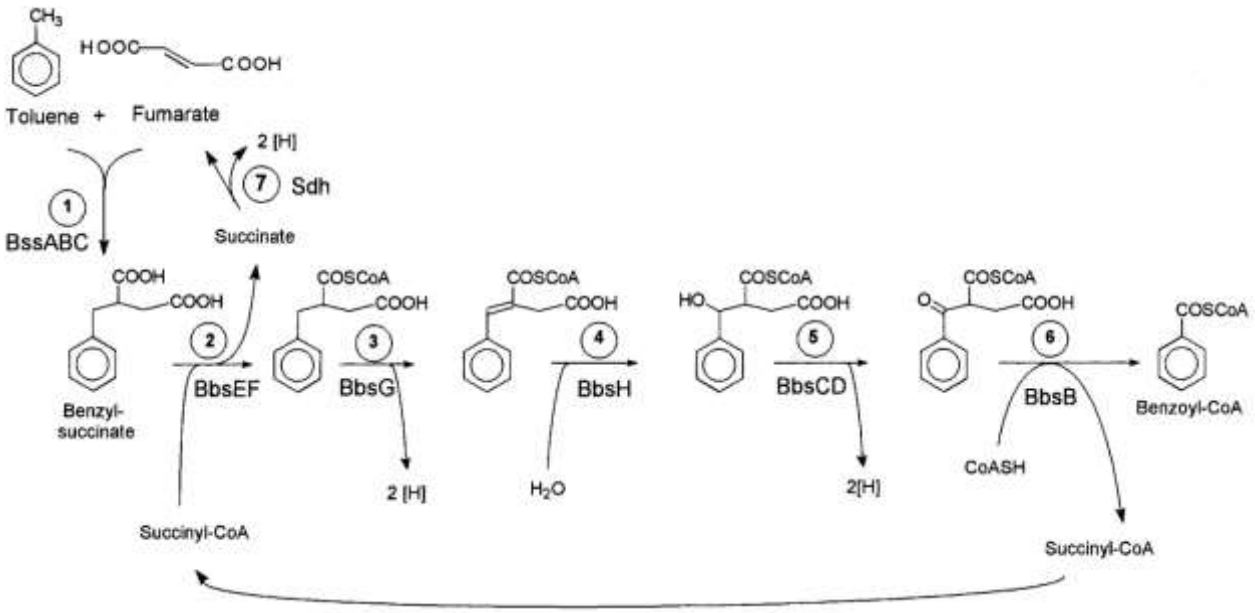


Figure 2.2.15. The catabolic pathway for anaerobic toluene degradation. (1) benzylsuccinate synthase, encoded by BssABC; (2) succinyl-CoA:benzylsuccinate CoA-transferase, encoded by BbsEF; (3) benzylsuccinyl-CoA dehydrogenase, encoded by BbsG; (4) phenylitaconyl-CoA hydratase, encoded by BbsH; (5) 3-hydroxyacyl-CoA dehydrogenase, encoded by BbsCD; (6) benzoylsuccinyl-CoA thiolase, encoded by BbsB; (7) succinate dehydrogenase, encoded by *Sdh*. (Leuthner et al, 2000).

n-alkane and cycloalkanes are metabolised with similar strategy by sulfate-reducing microorganisms in which a fumarate addition step activates these substrates to yield alkyl-succinates and cycloalkylsuccinate derivatives, respectively (Musat et al., 2010, Rios-Hernandez et al., 2003). The metabolism of cyclohexane by a nitrate-reducing microorganism to cyclohexylsuccinate (Musat et al., 2010) and of ethylcyclopentane by a sulfate-reducing organism to cyclopentylsuccinate (Rios-Hernandez et al., 2003) have been reported before. In the case of the anaerobic degradation of hexadecane by this mechanism (Fig. 2.2.16), alkylsuccinate synthetase (encoded by *assABC*) binds a fumarate to the substrate to produce 1-methylpentadecyclisuccinate. This intermediate is converted by methyl-malonyl-CoA mutase (encoded y mcm) to 2-(2-methylhexadecyl)malonate, which is decarboxylated by the activity

of a carboxyl transferase to produce 4-methyloctadecanoate. This last product is directed into the beta-oxidation pathway for further catabolism (Callaghan et al., 2012a). In the genome of *D. alkenivorans* AK-01, these genes are located at two different loci (*assA1* and *assA2*), and there is no similarity between them (Callaghan et al., 2012a) (Fig. 2.2.17).

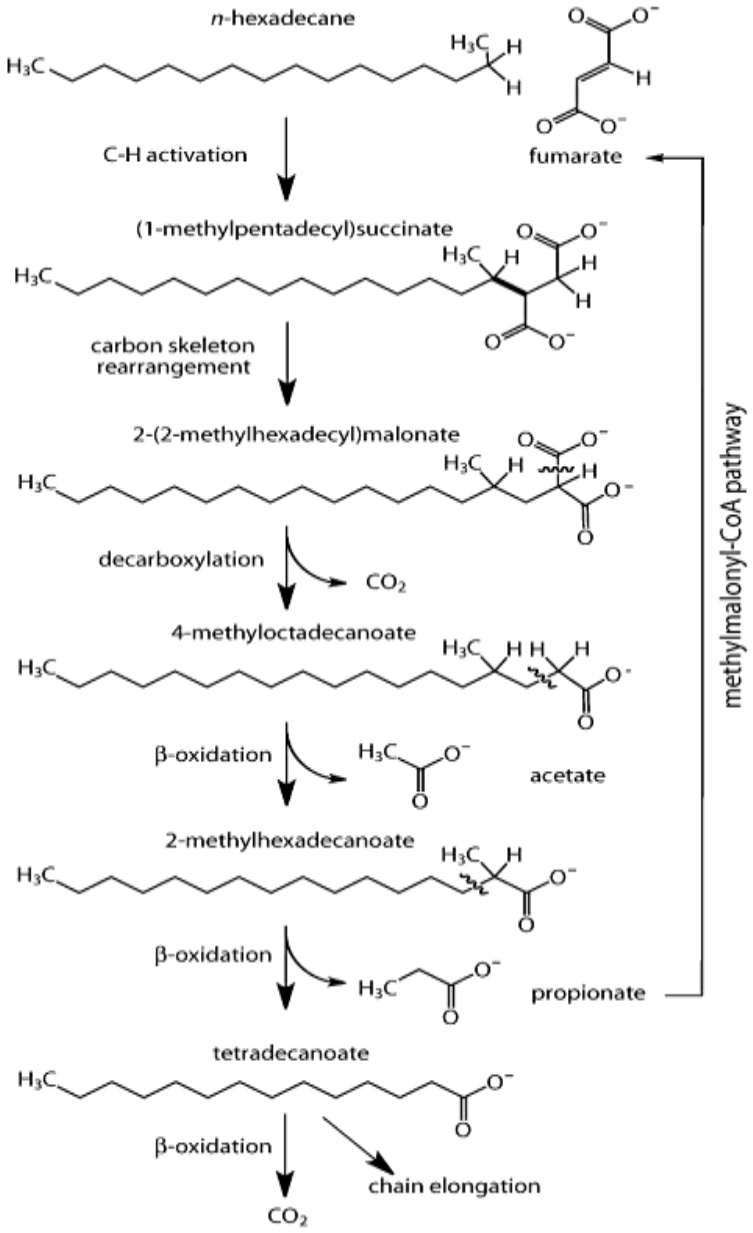


Figure 2.2.16. Biodegradation of *n*-hexadecane in *D. alkenivorans* AK-01 (Callaghan *et al, 2012).*

Several denitrifying bacteria use an oxygen independent hydroxylation process for the degradation of some aromatics in which a trimer (α, β and γ) molybdenum-containing ethylbenzene dehydrogenase (EBDH, encoded by *ebdABC*) hydroxylates the terminal carbon of this molecule to produce S-1-phenylethanol (Rabus et al., 2005, Szaleniec et al., 2007). The (S)-1- phenylethanol is converted into acetophenone and then to benzoyl-acetate and finally to benzoyl-acetyl-CoA as a result of reactions involving NAD-dependent (S)-1-phenylethanol dehydrogenase, acetophenone carboxylase (APC) and benzoylacetate-CoA ligase (BAL), respectively (Rabus et al., 2005, Szaleniec et al., 2007). The gene cluster for these three subunits of ethylbenzene dehydrogenase, along with the genes encoding (S)-1-phenylethanol dehydrogenase (*ped*) and a chaperone-like protein (encoded by *ebdD*) (necessary for transferring molybdenum into ethylbenzene dehydrogenase) are located on one operon, whereas the genes encoding APC (subunits A, B, C, D and E) and BAL are present on another operon (Rabus et al., 2005, Szaleniec et al., 2007).

Addition of a carboxyl group to substrates is an alternative reaction utilised for the anaerobic catabolism of hydrocarbons by several sulphate and nitrate reducing bacteria (Selesi and Meckenstock, 2009a). For instance, after conversion of propylene to acetone by *Xanthobacter autotrophicus* strain Py2, the acetone is carboxylated by acetone carboxylase at the expense of a CO_2 and an ATP to produce acetoacetate (Broberg and Clark, 2010). This multimeric ($\alpha 2\beta 2\gamma 2$) enzyme is encoded by three genes: *acxA*, *acxB* and *acxC* which encode the β, α and γ subunits, respectively, which are clustered in an operon as *acxABC* and its expression is regulated by a gene (*AcxR*) upstream of the *acxABC* cluster (Broberg and Clark, 2010).

$$CH_3\text{-}COCH_3 + CO_2 + ATP \rightarrow CH_3COCH_2COO + H^+ + AMP + 2Pi$$

Some microorganisms are able to anaerobically degrade alkenes and alkynes through the addition of H_2O to the unsaturated bond, producing the corresponding alcohols (Liu et al., 2011c). Using a monomeric thermostable acetylene hydratase (encoded by *AH* gene), for instance, *Pelobacter acetylenicus* adds a H_2O molecule to acetylene converting it to an acetaldehyde (Liu et al., 2011c, Tenbrink et al., 2011). In the degradation of β-myrcene (7-methyl-3-methylen-1,6-octadien) by *Castellaniella defragrans*, linalool dehydratase/isomerase (LDI, encoded by *ldi gene*) acts as a dual function enzyme that desaturates linalool to myrcene before isomerization to geraniol (Brodkorb et al., 2010a). The molecule is then oxidized by two dehydrogenases, referred to geraniol dehydrogenase (GeDH) *and* geranial dehydrogenase

(GaDH) (encoded by *geoA* and *geoB*, respectively) into geranial and geranic acid (Brodkorb et al., 2010a).

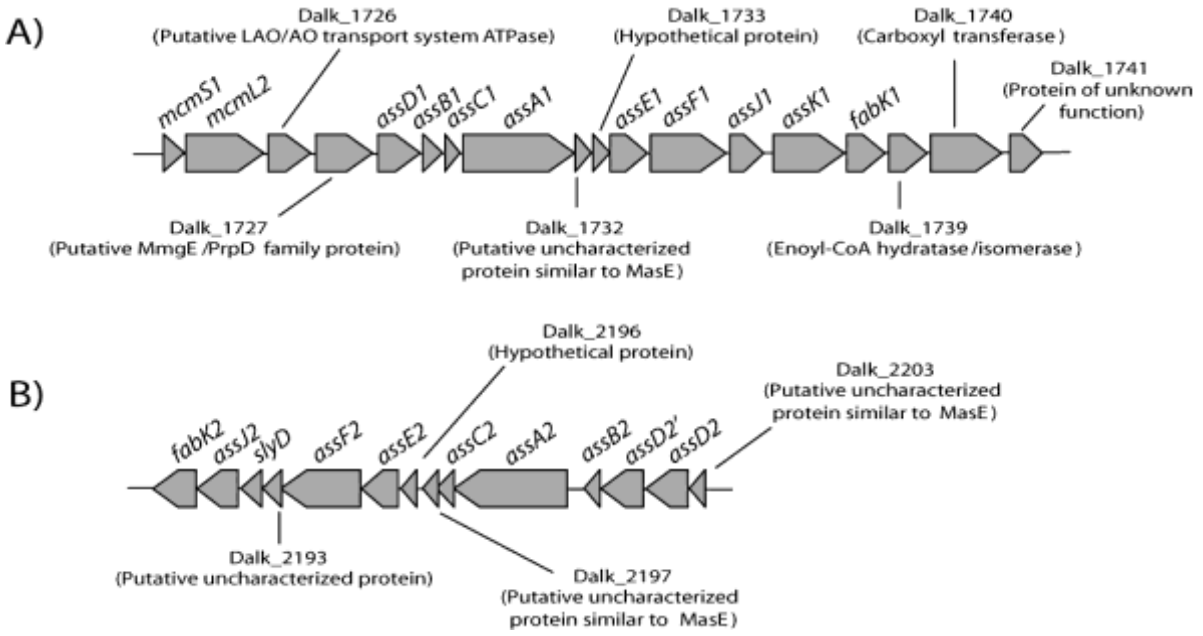


Figure 2.2.17. Gene organisation of the enzymes involved in hexadecane degradation in anaerobic conditions (in the *D. alkenivorans* AK-01 genome); A: Alkylsuccinate synthase locus 1 and B: locus 2; *mcmS1 and mcmS2:* small and large subunits of methylmalonyl-CoA mutase; *assD1*: alkylsuccinate synthase (I) glycyl radical activating enzyme; *assB1*, C1 and D1: alkylsuccinate synthase beta, gamma and alpha subunits; *assE1*: a chaperone protein; *assF1*: an uncharacterized protein; *assJ1*: enoyl-CoA hydratase/isomerase; *assK1*: AMP-dependent synthetase and ligase; *fabK1*: enoyl- (acyl-carrier protein) reductase II; *fabK2*: putative enoyl- (acyl-carrier protein) reductase II; *assJ2*: enoyl-CoA hydratase/isomerase; *slyD:* putative peptidyl-prolyl *cis-trans* isomerase (fkbp-type); *assF2*: putative uncharacterized protein; *assE2*: chaperone protein; *assC2, A2 and B2*: gamma, alpha and beta subunits (respectively) of alkylsuccinate synthase (II); *assD2′ and assD2*: alkylsuccinate synthase (II) glycyl radical activating enzyme (Callaghan *et al, 2012).*

Several methanotrophic microorganisms use methyl-coenzyme M reductase (Mcr) for the initial activation of methane, which then binds to coenzyme B (CoBSH) via methyl-coenzyme M (CoMSH) to produce a complex of CoM-S-S-CoB-heterodisulfide and methane (Thauer, 2011). This enzyme consists of two of each of the α, β and γ subunits, one nickel atom and a

tetrapyrrole cofactor (normally F_{430} factor or 17(2)-methylthio- F430) (Scheller et al., 2010, Thauer, 2011). The encoding genes for these subunits, *mcrBGA*, and two other additional genes with unidentified roles (*mcrC* and *mcrD*) are located on one operon (Scheller et al., 2010, Thauer, 2011). The *mcrA* gene is used as a marker to track methanogens and anaerobic methanotrophic microorganisms (Scheller et al., 2010, Thauer, 2011).

2.2.9 Anaerobic degradation of NACs

Several bacteria, such as sulfate-Reducing Bacteria and clostridia, are able to degrade nitroaromatic compounds in anaerobic conditions. The anaerobic metabolism of these compounds are performed through the reduction of the nitro group. However, it has to be mentioned that this mechanism is not exclusive to anaerobic bacteria and many other bacteria, such as *Enterobacter* sp., *Nocardiodes* sp. and *Rhodococcus* sp., are able to metabolise NACs using similar reductive pathways. The reductive pathways can be proceed by two different processes depending on the microbial gene capacity. *E. cloacae* PB2 expresses a monomeric flavoenzyme, referred to as PETN reductase (PETNr), which enables this microorganism to reduce triply nitrated aromatic compounds, such as picric acid and TNT. Furthermore, presence of *PnrA* gene, encoding for an NADPH dependent nitroreductase, enable several microorganisms to transform a variety of NACs, such as 3,5-dinitroaniline, 3- and 4-nitrobenzoate, 3-nitrotoluene, 2,4-DNT, TNT and 3,5-dinitrobenzamide.

Hydride transferases are a second reducing system in which the aromatic ring of some NACs, such as TNT and picric acid loss their nitrite. The initial reduction of TNT by PETNr, for instance, leads to production of hydride- and dihydride-Meisenheimer TNT complexes (H−-TNT and 2H−-TNT), which is further reduced through an unknown mechanism to release nitrite. *Nocardiodes simplex* FJ2–1A and *Rhodococcus erythropolis* are two bacteria with the ability to degrade picric acid using this mechanism. In these examples, a hydride is initially added to the aromatic ring by the activity of a F_{420}-dependent hydride transferase (*NpdI*) and F_{420} reductase (*NpdG*) to produce a hydride Meisenheimer complex. While in *N. simplex* the second hydride is added by the same enzyme, this step *R. erythropolis* is performed by the product of *NpdC*/*NpdG*. The dihydride Meisenheimer complex is later undergone a tautomerisation performed by the product of *NpdH* to release nitrite and various products.

2.2.10 Genetics of microbial adaptation to high hydrocarbon concentrations

A limitation of mineral nutrients is a common problem in oil contaminated marine environments due to an imbalanced C:P:N ratio (Megharaj et al., 2011). Microorganisms can adapt themselves to these limitations using either their own gene products or through the creation of a cooperative relationship with other microbes to decrease this stress (McCammick et al., 2010). Due to such adaptation processes, the community of hydrocarbon degrading microorganisms may show a sudden increases in cell mass after an initial temporary drop in the total number of microorganisms (Megharaj et al., 2011). In addition to the genes responsible for degradation of hydrocarbons, many bacteria adapt their physiology to the shortage of mineral nutrition through increased expression of their existing ion transporters (Yakimov et al., 2007, Schneiker et al., 2006). *Alcanivorax borkumensis* strain SK2, for instance, adapts itself to various mineral deficiencies through the induction of the genes responsible for different transport proteins such as *narKGHJI* cluster and *nrtCB–nasDTS* cluster (for reduction and uptake of nitrogen), *amt* (for uptake of ammonium), *phoBR* and *phoU–pstBACS* gene cluster (for uptake of phosphate), *znuAB* (for uptake of zinc), *modABC* (for uptake of molybdite), *mgtE* (for uptake of magnesium) and CorA-like MIT(for uptake of cobalt) (Schneiker et al., 2006).

In addition, one of the critical factors for biodegradation of hydrocarbons is the ability of the degrading microorganisms to be resistant to high concentrations of hydrocarbons, especially when the cells are suddenly exposed to large amounts of the compounds (Krajewski et al., 2014, Gallegos et al., 2008). A common resistance strategy to toxic solvents is to intensify the cell membrane density using isomerisation of Cis-unsaturated fatty acids to their Trans forms (McCammick et al., 2010). Furthermore, penetration of the solvents into cells induces expression of several chaperons to refold the proteins denatured by the reagents (Krajewski et al., 2014). Some bacteria tolerate high concentrations of solvents through expression of efflux pumps on their cell membrane, enabling them to export the toxic solvents out of the cells (Gallegos et al., 2008). *P. putida* S12, for instance, is a highly resistant strain to organic compounds due to the expression of a solvent resistance pump (encoded by *srpABC*), which discharges several types of solvents, hydrocarbons included (Sun et al., 2011).

Furthermore, many bacteria use chemotactic strategies to improve their resistance to different toxic compounds, hydrocarbons included (Krell et al., 2012, Fondi et al., 2013). The integral cell membrane associated proteins encoded by *alkN* (Fondi et al., 2013) and *nahY* in *P. putida* (Krell et al., 2012), for instance, interact respectively with alkanes and naphthalene. These

proteins act as a methyl-accepting chemotaxis factor that triggers a cell signalling pathway, regulating the flagella motor and resulting in cell attraction towards naphthalene (Krell et al., 2012, Fondi et al., 2013). Interestingly, both chemotactic activities and solvent resistance abilities are gene-dose dependent, and the amounts of the corresponding genes and the level of gene activity in cells determine the level of cell susceptibility to a solvent (Lacal et al., 2011). Based on a study, performed by *Lacal* (Lacal et al., 2011) on *P. putida* DOT-T1E, the presence of two alleles of chemoreceptor *mcpT* gene genes is enough to enable the cell to response strongly to different aromatic hydrocarbons. They showed that the level of *mcpT* gene expression in this strain was directly under control of the substrates, and there was an inverse relationship between the amounts of toluene and the level of gene methylation. The increases in the level of *McpT* methylation induce the activity of flagella motor using the autophosphorylation of *CheA* (Lacal et al., 2011).

The behaviour of bacteria in the case of exposure to a carbon source, which is potentially toxic for the cells, depends to the reaction of signal transduction proteins. Although the signalling pathway leads to triggering a degradation shunt in the cells, some bacteria prefer to migrate away from these compounds. This decision making in *P. putida* DOT-T1E in the case of exposure to toluene depends on the activity of *TodS*/*TodT* two-component system and *TtgV* in the cells. Since affinity of the sensor kinase *TodS* to toluene is double in comparison to *TodV*, this signalling protein is activated in lower concentrations of the substrate, inducing the TOD pathway (Busch et al., 2010). However, binding of toluene to *todV* in higher concentrations of toluene up-regulates expression of *TtgGHI* efflux pump to enhance the cell's resistance to the toxicity of this compound (Rodríguez-Herva et al., 2007).

2.2.11 Conclusion

Although there are some slight differentiations between many hydrocarbon degrading genes in different phylogenetic groups of microbial species, the homology of DNA sequences and organisation of these genes as well as intensive overlapping of the activity of their products indicates horizontal gene transfers have occurred between these groups. Slight differences in the corresponding hydrocarbon degrading genes can change the ability of microbes to degrade hydrocarbons, including altering the time of expression (due to presence of inducible or constitutive promoters), the level of expression (due to the activity of the promoter) and the enzymatic activity of the product (due to sequence of amino acids and the protein configuration).

This information assists environmental microbiologists and biotechnologists to choose suitable/stronger hydrocarbon degrading genes and the hydrocarbon resistance genes in order to create efficient genetic engineered microorganisms (GEM). Furthermore, such information can be used for the selection of appropriate (non-genetically engineered) microbial consortia with higher hydrocarbon degrading ability for use as an inoculum in the bioremediation of contaminated sites.

2.3 The Integration of Sequencing and Bioinformatics in Metagenomics

Published in: Reviews in Environmental Science and Bio/Technology; 14(3): 357-383.

2.3.1 Introduction

Microorganisms are ubiquitous in the environment and are found in soils, waters, plants and animals, as well as in extreme environments, including deep seas, volcanoes and frozen habitats (Lewin et al., 2012, Stefanis et al., 2013). The cultivation of microbes and microscopic enumeration have been the traditional methods used for microbial characterization (Stefanis et al., 2013). However, traditional cultivation techniques are unable to identify the majority of microorganisms in environmental samples due to the inability of many organisms to grow on laboratory media (Mocali and Benedetti, 2010). Indeed, laboratory cultivation techniques are estimated to support the growth of only 1% of the microorganisms in a bacterial community so the majority of the microbes are missed using these methods (Vieites et al., 2009, Dionisi et al., 2012). However, DNA sequencing techniques, especially those targeting ribosomal RNA genes, have opened new windows for investigating uncultured bacteria in different environments (Joshi et al., 2014). Sequencing of PCR products derived from 16SrRNA genes and 5SrRNA genes from different communities provides newer culture-independent procedures to reconstruct the phylogeny of microorganisms and to investigate the diversity and abundance of the microorganisms in a community (Rajendhran and Gunasekaran, 2011).

Metagenomics is a technique in which the genome contents of a whole community of organisms, especially microorganisms that inhabit a specific niche, is sampled in order to investigate the community structure and also to screen for special traits or DNA sequences (Desai et al., 2012). The term “metagenomics” is derived from the combination of a huge statistical process for analysing data (meta-analysis and the widespread and whole genomic analysis of a community of organisms (genomics) (Schloss and Handelsman, 2003). Many advances in soil and aquatic microbial ecology, including studies on single genes, complex pathways, microbial diversity, and the role of specific microbes in a desired community, have resulted from progress in sequencing technology, the polymerase chain reaction (PCR) procedure, and gene cloning strategies (Simon and Daniel, 2011) . In principle, a metagenomic approach is applicable to any environment (such as soil, aquatic habitats, hospital environment)

where microbial genomes can be obtained from samples. One of the most interesting metagenomic studies, the Human Microbiome Project, is mapping all the microbial communities that are associated with the human body (Aagaard et al., 2013). In this paper, the methodology and some applications of metagenomics is discussed.

2.3.2 Technical basis of metagenomics

Technically, after DNA extraction, the genomic material can be analysed via two ways. In sequencing independent techniques, the biodiversity of particular amplicons (the PCR-amplified sequences) can be estimated by microarray or a PCR-dependent fingerprinting technique, e.g. RFLP (Restriction Fragment Length Polymorphism) (Andoh et al., 2011), T-RFLP (Terminal-Restriction Fragment Length Polymorphism) (Sjöberg et al., 2013), DGGE (Denaturing Gradient Gel Electrophoresis) (Green et al., 2010), AFLP (Amplified fragment length polymorphism) (Hoffman et al., 2012), RAPD (Random Amplified Polymorphic DNA) (Katara et al., 2013) and SSCP (Single-strand conformation polymorphism analysis) (Schmalenberger and Tebbe, 2014).

However, metagenomics relies on sequencing of the genome and comparing the sequence of fragments of interest with known sequences (Huang et al., 2013). These studies can be adapted for analysis of both total extracted DNA and specified (functional or systematic) DNA molecules in a sample (Huang et al., 2013). When RNAs are analysed in metagenomic studies, the technique is termed “metatranscriptomics” (Gosalbes et al., 2011). The sequence-driven analysis employs conserved DNA sequences as a template for designing primers to screen for a desired sequence in environmental DNA samples or a metagenomic library (Shah et al., 2013). Based on phylogenetic anchors, including 16S rRNA in bacteria and radA gene (DNA repair gene) in archaea, the technique is able to provide phylogenetic information on the organism from which the sequences (clones) were derived (Singh et al., 2009). Furthermore, whole genomic DNA sequencing, in addition to using the 16S rRNA gene sequences, provides considerable information regarding the role of the organisms in their community (Su et al., 2012). In the other word, determination of the quantity of critical functional genes, such as dissimilatory sulphite reductase, hydrogenase and nitrate reductase genes, can indicate the biological activities occurring in an environment (Su et al., 2012).

The existing sequencing systems can be divided into first generation (Sanger), second generation (including pyrosequencing, Illumina and SOLiD) and third generation (including

HeliScope, PacBio RS, Ion Torrent and GridION) (Shokralla et al., 2012). The technical basis of these techniques and their limitations and advantages are summarized in table 2.3.1. The chain terminator (Sanger) based sequencing approach was the first technique used for the sequencing microbial genomes in which DNA polymerase uses a single strand DNA as a template and incorporates nucleotide (dNTPs) from the 3/ terminus of primer (Ansorge, 2009, Shendure et al., 2011). A dideoxynucleotide (ddNTP) pool is prepared using a mixture of both dideoxynucleotides (ddNTPs) from each base labelled with different fluorophores and their normal dNTPs (deoxynucleotides). Random incorporation of the fluorescently labelled dideoxynucleotides (ddNTPs) terminates each primer extension reaction, generating many end-fluorescent labelled extension products which (based on their molecular weight) are separated from each other by capillary electrophoresis with the fluorophore which terminates each band being identified under UV light (Ansorge, 2009, Shendure et al., 2011). This approach is able to determine DNA sequences of >700 bp with an accuracy of 99.999%, and at a cost of 0.50 USD per Kb (Shendure and Ji, 2008, Thomas et al., 2012). In traditional random "shotgun" sequencing using this form of Sanger DNA sequencing (dye terminator sequencing), the DNA fragments are cloned into plasmid vectors with suitable priming sites for sequencing either side of the cloning site (Jianping, 2006, Shendure et al., 2011). Each clone has to be individually sequenced, which limits the speed of the technique, although the read lengths may be up to a kilo base (Jianping, 2006, Shendure et al., 2011). In addition to this limitation, high demands of genome sequencing projects and metagenomics have led to the development of next generation sequencing technologies (Ansorge, 2009, Shendure et al., 2011).

However, the cyclic-array-based techniques, including the Illumina/Solexa, the Roche/454 GS FLX Titanium sequencer), the HeliScope Single Molecule Sequencer technology, the Polonator and the Applied Biosystems (SOLiD), work based on fragmentation of the DNA sample and ligation of the fragments to common adaptor sequences that are then used as priming sites for PCR based amplification procedures (Masoudi-Nejad et al., 2013a, Koboldt et al., 2013). These approaches work by sequencing a compact array of genomic fragments using continuous cycles of enzymatic reactions and collection of the data based on released photons (or ions in the case of ion torrent machines) (Metzker, 2009, Koboldt et al., 2013). Based on these methods, millions of sequencing reactions are carried out simultaneously in parallel as an array on a chip. Here two of the most used cyclic array sequencing methods (the Illumina/Solexa, the Roche/454 GS FLX Titanium sequencer and the Illumina/Solexa) are discussed in more details (Metzker, 2009, Koboldt et al., 2013).

The Roche/454 GS FLX Titanium sequencer (based on the pyrosequencing approach) is a real time method that utilises a combination of enzymatic reactions to synthesise DNA from templates, with a different nucleotide supplied at each cycle, and simultaneous bioluminescence to monitor the amount of pyrophosphate released at each cycle from each template (Fakruddin et al., 2013). ATP-sulfurylase activity converts the pyrophosphates released through the polymerization reactions to ATP. The ATP provides the energy required for luciferase to produce visible light that is recorded by a LCC camera. The extra-nucleotides are eliminated by Apyrase to prepare the media for addition of a second nucleotide. Based on the linear correlation between the light emitted and the nucleotides incorporated this technique can also determine how many nucleotides are added at each cycle to a template, and therefore can correctly sequence regions with several identical bases in a row. (Royo et al., 2007, Fakruddin et al., 2013). In practice, pyrosequencing is automatically performed by carrying out PCR amplification of amplicons (target DNA ligated to adaptors)-bearing beads (where each bead has one molecule of an amplicon) with *Bst* polymerase (from *Bacillus stearothermophilu)* and SSBPs (single-stranded binding proteins) in order to provide enough template for the sequencing reaction, and then transferring each bead in a separate well placed on a micro-fabricated plate (Fakruddin et al., 2013, Delseny et al., 2010). Other necessary enzymes, such as luciferase and ATP sulfurylase, are also immobilized on the beads (Fakruddin et al., 2013, Delseny et al., 2010). Throughout the sequencing process the necessary agents are added to the wells and the photons emitted by luciferase are read by a charge coupled device (CCD). Alternatively, the phosphate ions released from pyrophosphate can be measured directly in ion torrent devices (Fakruddin et al., 2013, Delseny et al., 2010).

Table 2.3.1 Comparison of three generations of DNA sequencing methods. * GridIon is sometime referred to as fourth generation of sequencers

Classification	Platform Name	Sequencing method	Method of Detection	Analysis time	Length (bp)	Raw error rate (%)	Data output	Advantages	Limitations	References
First generation	Sanger 3730xL (designed by ABI/Life technologies company)	Sanger	Fluorescence/Optical	48 hours	Up to 900	0.001	1.9~84 Kb	Long read lengths; high accuracy	Low throughput; high cost;	(Liu et al., 2012; Solomon et al., 2014)
Second generation	Genome Sequencer FLX+ System (designed by Roche/454 company)	emulsion PCR or Pyrosequencing	Photons detected by charge coupled device	23 hours	500–600; Up to Kb	1	700 Mb	Long read; high throughput; 1–5 µg DNA needed	Expensive instrument; difficult sample preparation; high errors	(Liu et al., 2012; Solomon et al., 2014)
	HiSeq 2000 (Designed by Illumina company)	synthesis by reversible terminator sequencing	Fluorescence/Optical	9–15 days	101–151 bp	<1	95-600 Gb	Very high throughput; <1 µg DNA needed	Short length reads; Expensive instrument; high costs; high errors; slow runs	(Liu et al., 2012; Solomon et al., 2014)
	MiSeq (Designed by Illumina company)	synthesis by reversible	Fluorescence/ Optical	5 hours	25-300 bp		13-15 Gb	Very high throughput; Fast	Short read length	(Liu et al., 2012; Solomon et al., 2014)
	5500xl SOLiD System (designed by ABI/SOLiD company)	Base ligation	Fluorescence/Optical	7 days	85 bp	0.002	120 Gb	Very high throughput; low errors; 2–20 µg DNA needed	Expensive instruments; short reads; high costs; slow runs	(Liu et al., 2012; Solomon et al., 2014)

Third generation	HeliScope (designed by Helicos company)	Single-molecule sequencing	Fluorescence/Optical	8 days	200	> 2	21–35 Gb	High throughput than second generation	Expensive instruments; Short reads; high costs	(Solomon et al., 2014)
	PacBio RS (designed by Pacific Biosciences company)	Real-time single molecule sequencing	Fluorescence/Optical	30-120 min	2500-23000 (average 2246)	1.3	54.5 Gb	High throughput than second generation; long read; real-time detection	expensive instrument; inactivation of polymerase in ZMWs; high cost per base	(Au et al., 2012; Ferrarini, 2012; Koren et al., 2012; Liu et al., 2012; Solomon et al., 2014)
	Personal Genome Machine (PGM) Sequencer (designed by Ion Torrent/Life Technologies)	Sequencing by synthesis	Changes in proton concentration (pH)	3.5 h	>400	0.3	1.3 Gb in case of using a	Fast; long reads; no need to conjugated bases	potential errors in case of homopolymers and highly repetitive sequences	(Butler et al., 2012; Merriman et al., 2012; Solomon et al., 2014)
	GridION* (designed by Technologies and Researchers at the University of Oxford)	Nucleotide scanning by nanopores	Current	15 minutes	< 10kb	------	------	Very fast (1 bp/10 ns); low costs; Lower errors; Whole-human genome sequencing in only 15 min	Not commercially available	(Eisenstein, 2012; Maitra et al., 2012)

The Illumina/Solexa method is performed by the Illumina Genome Analyzer (GA) and is based on sequencing-by-synthesis (Liu et al., 2012). First, both ends of the DNA fragments are linked to adaptors and the DNA molecules are denaturized. The single strand fragments are then added to flow cell channels where they can randomly bind to polymer-coated glass surfaces. Subsequently, an annealing step binds sequencing primers to the DNA. Each dNTP used during the subsequent sequencing procedure in this method is bound to a fluorophore so that, DNA

polymerization is terminated after incorporation of the first nucleotide. The enzyme and unused nucleotides are removed from the cells and the incorporated base is identified based on colour of its label and using two lasers (green: G, T and red: A, C) and different filters. Next, the fluorophore is removed from the integrated base allowing the second round of polymerization. These cycles then continue for the length of the sequencing run (Park et al., 2013, Kircher and Kelso, 2010).

The cyclic array sequencing of signature sequences such as regions of 16S rRNA genes (usually pre-amplified by PCR using primers specific for conserved sequences) is used for the determination of microbial diversity in the environment (Caporaso et al., 2011, Solomon et al., 2014). Since the length of sequenced fragments are quite short (up to 500bp by Roche 454 GS FLX+ and less than 200 bp by Illumina HiSeq 2000) (Shokralla et al., 2012, Luciani et al., 2012), application of these systems for microbial phylogenetic studies is limited to hyper-variable regions of the marker genes, such as 16S rRNA (Soergel et al., 2012). On the other hand, because these approaches are able to produce thousands of sequence reads for a given sample, they allow identification of all of the components of a mixed community of microorganisms (Ye and Zhang, 2013). Since all sequencing processes, including construction of a library of amplicons and amplification, are performed *in vitro* using PCR-based techniques, the limitations of transformation of a strain, colony isolation and the need to purify individual genes for sequencing are avoided (Tamaki et al., 2011). Furthermore, since all the individual sequencing templates are fixed in place on a planar surface and the reagents are added and removed as necessary as a single volume of reagent, application of array-based sequencing methods enhances massively the rate of sequencing (Stranneheim and Lundeberg, 2012, Church et al., 2011). Finally, the ability to produce far more sequence data compensates for the shorter read lengths of this technology making the cyclic-array-based technologies much more cost efficient than Sanger sequencing strategy for large sequencing projects (Stranneheim and Lundeberg, 2012, Church et al., 2011). At present, however, most of these technologies are only used for biological researches as their high costs and low throughput limit their application for routine medical tests (Pickrell et al., 2012, Masoudi-Nejad et al., 2013b). Much work has recently been conducted to develop these approaches from small scaled research applications to routine laboratory tests with the ability to examine tens to hundreds of millions of genomes per year, and applicable even for routine diagnostic tests in medical laboratories. These researches have led to the emergence of alternative DNA sequencing methods, referred to as third generation of sequencing approaches.

The Ion Torrent (designed by Life Technologies), for instance, is one of the third generation of sequencing methods which uses sequencing-by-synthesis during which the release of pyrophosphate through incorporation of a matching base extending the DNA changes the pH of the media. This small change in the media is detected by pHFET (pH-sensitive field effect transistor) and the data is processed by CMOS (complementary metal-oxide-semiconductor) chip. The four bases are added in sequence to the media and the unmatched bases are washed out from the media. As in pyrosequencing systems, homopolymeric sequences are indicated by a proportionate increase in pH value (Merriman et al., 2012).

Single-molecule sequencing (SMS) systems such as the Helicos Biosciences true Single Molecule Sequencing (tSMS) and Pacific Biosciences Single Molecule Real-Time (SMRT™) sequencing use a sequencing-by-synthesis technique in which a fluorescently labelled nucleotide and DNA polymerase is repetitively flowed through the cells in order to synthesise the template of a given single strand DNA (Shokralla et al., 2012, Radford et al., 2012). In the tSMS platform, due to the anchoring of fluorescent compounds to the nucleotides, incorporation of each nucleotide into the extending DNA blocks DNA synthesis until the residue is released. Release of this compound is associated with release of a fluorescence which is detected by a CCD camera. The Next round of nucleotide incorporation is followed by washing the cell and removing the remnants (Radford et al., 2012). The PacBio RS is a single-molecule real-time sequencer (SMRT) in which a nano-structure referred to as Zero Mode Waveguide (ZMW) record continuously DNA polymerization. ZMW is a nanostructure made of tens of thousands of holes with sub-wavelength, 10 nanometre, diameter length where thousands of DNA polymerization reactions are performed in parallel. The DNA polymerization from the template is performed by a DNA polymerase located at the bottom of each pore. Since the fluorescence moiety is linked to the phosphate of nucleotides, each nucleotide incorporation is associated with light emission. Removal of the washing step from the system accelerate its rate of nucleotide incorporation and thereby improve the sequencing quality. This system is faster than tSMS and provide longer read lengths with lower error rates (Ku and Roukos, 2013).

Nanopore equipped gene sequencer, referred to as fourth generation of sequencers, works through monitoring ionic currents during the translocation of a single strand of DNA through a nanopore (Maitra et al., 2012). Nanopores are natural or engineered nano-scale channels

installed into a membrane. The biological pores are created by the installation of a protein channel, such as Staphylococcal α-haemolysin (Masoudi-Nejad et al., 2013b) and MspA (*Mycobacterium smegmatis* porin A) (Manrao et al., 2012) into a lipid membrane while the solid-state pores are produced via etching or drilling a channel in a solid-state layer (dela Torre et al., 2012). While a single strand DNA (ssDNA) is passing through the pore, a voltage is applied across the basement and the ionic current changes through the channel, caused by the electrostatic forces of the nucleotides, is monitored by a patch-clamp amplifier. Since this platform works only by monitoring an electrical current and is sensitive enough to track small amounts of DNA in a sample, there are no needs for sample preparation, reagents, replicative enzymes and detector optics (dela Torre et al., 2012). The translocation rate of nucleotides through pores is very fast (>1nt/10 μs for the biological pores and >1nt/10ns for the solid-state pores), so this system is potentially able to distinguish nucleotides at speed (Timp et al., 2010). Furthermore, this system is potentially able to improve the read-length of contigs to more than 10kb, which in turn, decrease the assembly error and costs.(Maitra et al., 2012). The GridION is a high-throughput nanopore-based sequencer introduced by Technologies and Researchers at the University of Oxford, in which an array of 2000 separate nanopores enables researchers to sequence human genome in only15 min (Han et al., 2013, Pickrell et al., 2012). Also, the MinION, designed by the same group, is a "USB-sized" nanopore-based sequencer with 512 nanopores with similar qualifications (Pickrell et al., 2012, Henson et al., 2012).

To sum up, selection of one of these platforms depends deeply on the purpose of the analysis. Since the Roche/454 GS FLX machine produces long length reads, and accurate identification of a strain based on 16SrRNA need reads with more than 400 nt, this approach is a suitable approach for biodiversity studies. However, HiSeq Illumina is recommended for full metagenomic analysis since it produce a huge amounts of data per sample.

2.3.3. Bioinformatics and Metagenomics

Reconstruction of genome libraries from metagenomic data is an approach useful for both sequence information and the identification and classification of novel genes or new found metabolic pathways (Davenport and Tümmler, 2013). However, the huge amounts of data obtained from metagenomic studies provides researchers with a computational challenge that is resolvable only by powerful bioinformatics tools (Davenport and Tümmler, 2013). Bioinformatic tools are now powerful enough to predict the genomic structure of organisms and the probable function of specific genes (Abbai et al., 2012). This ability is of interest to

many researchers who work on comparative genomics and metagenomics. All algorithms are based on four aspects of metagenomics : 1) The given sample is very complicated in terms of the existing microorganisms, 2) each sample contains different ratios of each constituent species, 3) millions of reads must be analysed and 4) horizontal gene transfers lead to some genes being shared between different microorganisms (Guo et al., 2013).

These software are able to analyse the given sequence and to compare the information with the existing data in several international collaborative gene databases **(**a list of most useful gene databases is listed in table 2.3.2. Many of these databases are currently linked together in order to improve the accuracy of sequence analysis. Gene3D (http://gene3d.biochem.ucl.ac.uk), for instance is a protein domain database that is integrated with SUPERFAMILY and Pfam to offer a very comprehensive protein domain database (Lees et al., 2014). The Gene3D itself, in addition to DomSerf, FUGUE, pDomTHREADER and Phyre, supports another web-based program referred to as Genome3D (http://www.genome3d.eu) which indicate 3D structure of a protein and the accurate position of its active site, and so it is able to determine the relationships between sequence of a protein and its structure and function (Lewis et al., 2013). Such integrated databases need a powerful retrieval system which is able to connect to all the information in each database without changing any information. Sequence retrieval system (SRS) at EBI and Entrez at NCBI and many other centralizing systems are some of the integrating systems that are able to access data deposited in more than 250 databases and to link between different databases (Pannarale et al., 2012, Luo, 2013, Acland et al., 2013).

Source	Full name	Specified for:	References
GenBank	Gene Bank	An annotated collection of all publicly available DNA and protein sequences present in GenBank at NCBI, DNA DataBank of Japan (DDBJ) and the European Molecular Biology Laboratory (EMBL). These three organizations exchange data on a daily basis.	(Mizrachi, 2013
DDBJ	DNA DataBank of Japan	nucleotide sequence data	(Kosuge et al., 2014
EMBLEBI	European Molecular Biology Laboratory-European Bioinformatic Institute	A comprehensive database for raw and assembled sequences	McWilliam et al., 2013
RDP	Ribosomal Database Project	It includes bacterial and archaeal small subunit rRNA genes as well as fungal large subunit rRNA genes	Chai and Cole, 2014
SILVA	From Latin silva	A datasets of aligned small (16S/18S, SSU) and large subunit (23S/28S, LSU) ribosomal RNA (rRNA) sequences for *Bacteria***,** *Archaea* and *Eukarya*.	Quast et al., 2013
GOLD	Genomes Online Database	A comprehensive resource for global monitoring of genomic and metagenomic studies.	Pagani et al., 2012
EuPathDB	The Eukaryote Pathogen DataBase	This database include 11 databases to support eukaryotic infectious disease pathogens and invertebrate vectors of human disease.	
UNITE	User-friendly Nordic ITS Ectomycorrhiza Database	A rDNA sequence database designed identification of ectomycorrhizal asco- and basidiomycetes.	Abarenkov et al., 2010
fungiDB	Fungi DataBase	It contains the genome sequence and annotation as well as cell cycle microarray data, hyphal growth RNA-sequence data and yeast two hybrid interaction data obtained from 18 species spanning several fungal classes, including the Ascomycota classes, Basidiomycota, *Saccharomycetes Sordariomycetes and* orders, *Tremellomycetes and Pucciniomycetes*, and the basal *'*Zygomycete' lineage *Mucormycotina*.	Stajich et al., 2012

VIPR	Virus Pathogen Database and Analysis Resource	A database to serve sequence records, gene annotations, protein annotation, immune epitope locations, protein conformation, clinical and surveillance metadata and novel data derived from comparative genomics analysis obtained from several human pathogenic viruses including the Bunyaviridae, Caliciviridae, Coronaviridae, Flaviviridae, Filoviridae, Hepeviridae, Herpesviridae, Paramyxoviridae, Picornaviridae, Poxviridae, Reoviridae, Rhabdoviridae, Togaviridae and Arenaviridae families.	Abarenkov et al., 2010
JGI	Joint Genome Institute	A database for environmental and evolutionary microbiology, especially for the microorganisms resistant to extreme conditions such metal contamination, radiation and high acidity.	Grigoriev et al., 2012
Rfam	The RNA Family database	a collection of RNA families (non-coding RNA genes, structured cis-regulatory elements and self-splicing RNAs), each represented by multiple sequence alignments, consensus secondary structures and covariance models (CMs)	Burge et al., 2013
rCAD	The RNA Comparative Analysis Database	It employs three different dimensions of information (sequence of nucleotides, 2-D structure and evolutionary relationships) for RNA comparative analysis datasets	Ozer et al., 2011
UniPort	Universal Protein Resource	A comprehensive resource for protein sequence and annotation data integrated by EMBL-EBI, the SwissPort and Protein information resource.	Wu et al., 2006
SwissPort	Swiss Protein Resource	A database for structure, domain and function of proteins	Bairoch and Apweiler, 2000
SUPERFAMILY	SUPERFAMILY	A database for protein domains and structure	de Lima Morais et al., 2011
Pfam	The protein family database	A database for protein-protein interaction	Finn et al., 2014
Gene3D	Gene three dimensional configuration	integrated with SUPERFAMILY and Pfam to offer a very comprehensive protein domain database	(Lees et al., 2014)
Genome3D	Genome three dimensional configuration	An integration of Gene3D, DomSerf, FUGUE, pDomTHREADER, and Phyre which indicate 3D structure of a protein and the accurate position of its active site, and so it is able to determine the relationships between sequence of a protein and its structure and function	(Lewis et al., 2013)
BioGRID	The Biological General Repository for Interaction Datasets	A database for protein-protein interaction	Chatr-aryamontri et al., 2013

MINT	The Molecular INTeraction database	A database for protein-protein interaction	Licata et al., 2012
RegulonDB	Regulon DataBase	A database for transcription factors	Salgado et al., 2013
PDBe	Protein Data Bank in Europe	Three-dimensional structures of proteins	Gutmanas et al., 2014
PDBj	Protein Data Bank Japan	Three-dimensional structures of proteins	Kinjo et al., 2012
DOMMINO	Database Of Macro-Molecular INteractiOns	A fully automated database to cover different types of molecular interactions, domain–domain interactions, domain–peptide interaction, integrated to SUPERFAMILY.	(Kuang et al., 2012
TRANSFAC	TRANScriptional FACtor	A database for transcriptional factor and their DNA binding sites	Wingender et al., 1996
KEGG	Kyoto Encyclopedia of Genes and Genomes	A database for metabolic pathways	Kanehisa et al., 2012
MANET	The Molecular Ancestry Network	A database for metabolic pathways with the ability to map evolutionary relationships of protein architectures onto biological networks	Kim et al., 2006
EcoCyc	*E. coli* encyclopaedia	A database for metabolic pathways using *Escherichia coli* K-12 MG1655 as a source microorganisms. The EcoCyc project performs literature-based curation of the entire genome, and of transcriptional regulation, transporters, and metabolic pathways.	Keseler et al., 2013
MetaCyc	Metabolic encyclopaedia	A database of non-redundant for metabolic pathways, reactions, enzymes, genes and chemical compounds. Several quantitative data, including enzyme kinetics, are included in this databae	Caspi et al., 2014

Table 2.3.2. A list of Gene database and their features.

Before analysis in these systems, the millions of short reads obtained from the sequencing step need to go through several filtration processes, referred to as flowgram clustering or denoising,

to remove barcodes, and primers as well as non-informative and low quality sequences. Sickle (https://github.com/najoshi/sickle), SolexaQA (Cox et al., 2010), Trimmomatic (http://www.usadellab.org/cms/?page=trimmomatic), PRINSEQ (http://prinseq.sourceforge.net/), FASTX quality trimmer (http://hannonlab.cshl.edu/fastx_toolkit/), ERNE-FILTER (http://erne.sourceforge.net/), CONDeTri (Smeds and Künstner, 2011), Kraken (Kong, 2011), Cutadapt (Martin and Rahmann, 2012), AlienTrimmer (Criscuolo and Brisse, 2013), NextClip (Leggett et al., 2013), StreamingTrim 1.0 (Bacci et al., 2013) and ngsShoRT (Chen et al., 2013) are some of the most famous tools for quality control of the reads. Although their ability is different, most of them are able to trim adaptor sequences and low quality ends and to remove the reads containing ambiguous nucleotides, duplicate reads and high masked reads from the sequencing data output to accelerate data read assembly and read data analysis (Del Fabbro et al., 2013, Kong, 2011). The read assemblers are categorised into two groups: 1) the de Bruijn graph based approach, such as Velvet (Zerbino and Birney, 2008), SOAPdenovo (Li et al., 2010), PASHA (Liu et al., 2011d) and SPA (Yang and Yooseph, 2013) and 2) the overlap-layout strategy such as SHARCGS (Dohm et al., 2007), SSAKE (Warren et al., 2007) and PE-assembler (Ariyaratne and Sung, 2011). However, due to the production of too small contigs, these systems are not able to completely assemble the whole genome of an organism (Ramos et al., 2013). In other words, since in these approaches the reads are assembled based on sufficient overlaps between short fragments there is a higher chance for mis-assemblies of the fragments (Bonham-Carter et al., 2013). Therefore, a hybrid approach such as Graphic Contig Analyzer for All Sequencing Platforms G4ALL (Ramos et al., 2013), MaSuRCA (Zimin et al., 2013) and PRICE (paired-read iterative contig extension) (Ruby et al., 2013) increase the sensitivity and accuracy of gene assembly. Although there are several softwares for trimming and assembly of the contigs, since this step is usually performed by the sequencer company or the metagenomic analyser pipeline, there is no pressure on the operators to select a right one for their own studies. After assembly of the contigs to full length DNA, their information is then analysed by phylogenetic and function-based software.

Selection of a bioinformatic tool for sequencing programs depends partly on the read length (Rachamalla et al., 2012). The chosen tool must be able to arrange all the individual metagenomic sequences (reads) without any error in order to determine the taxonomic diversity pattern of a given habitat; this is called binning (Rachamalla et al., 2012). However, most of the existing metagenomics sequencing recovery systems are undersupplied in terms of handling

the amount of data and the sensitivity and quality with which they assemble the contigs (Guo et al., 2013). To date, several software have been designed for handling such large amounts of data that can be classified into two categories: composition-based and alignment (comparative)-based software (Table 2.3.3) (Higashi et al., 2012, Katoh and Standley, 2013).

Purpose	Strategy	Name of software	Principle	Advantages	Disadvantages
Phylogenetic analysis	alignment-based methods	UniFrac, DOTUR, PHYLIP 3.66, MEGA4, DIALIGN TX, MOTHUR, LIBSHUFF, MrBayes 3.2, TreeOTU, DAMBE5, PhyloTreePruner, BlastGraph, LSCluster	These software match metagenomic sequences based on analysis of the small- subunit rRNAs (ssu-rRNAs) sequences in presence BLAST	-	1) Focus on the sequence homology 2) they are not perfect due to horizontal gene transfer phenomenon 3) expensive
	Composition-based methods	TETRA, Phylopythia, SSBA, MetaCV	using oligonucleotide frequencies of the genome, GC content and codon usage	Faster and lower cost than the alignment based method, they are less accurate	Short sequences do not have insufficient signals for taxonomic investigations
	Mix Strategy	A combination of alignment based methods and composition based methods	SPHINX, MetaSAMS and phiSpy	Fast and more accurate	-
Function analysis	homology-based programs	PipMaker, SGP-1, AGenDA, GENEWISE, BLANNOTATOR	using gene sequence databases such as Smith-Waterman algorithms and BLAST for prediction of the closest homologous sequence	-	limited to known genes
	splicing site prediction softwares	NetGene2, Genesplicer, ESEfinder, HSF	maximal dependence decomposition (MDD)	Prediction of genes	

	the Ab initio softwares	GENIE, Gene-Mark, GeneBuilder, GLIMMER, GENSCAN, ASPic-GeneID, MetaGeneTack, MESSA and FFPred 2.0	using consensus sequences, secondary structures, transmembrane helices, signal peptides, 3D structure and structurally disordered regions	Prediction of genes	Many genes are ignored because of short length of fragments
	Combination	MetaGene, MEGAN 5, MEta Genome Analyser, MG-RAST	A combination of other strategies	Mostly available online and free of cost	-

Table 2.3.3. List of bioinformatic software and their abilities.

The alignment-based methods are based on the BLAST algorithm and, match metagenomic sequences with existing databases to find the taxonomic group of the given sample and to construct phylogenetic trees of the sequences (Pearson, 2014, Bazinet and Cummings, 2012). These trees are based on analysis of the small- subunit rRNAs (ssu-rRNAs) sequences and show the sequence diversities of a habitat and determine the homology of amplified sequences between the inhabitants of the environment in evolutionary terms (Wu et al., 2013, Sunagawa et al., 2013). Up to now, several software such as, UniFrac (Lozupone and Knight, 2005), DOTUR (Schloss and Handelsman, 2005), PHYLIP 3.66 (Felsenstein, 2006), MEGA4 (Felsenstein, 2006), DIALIGN TX (Subramanian et al., 2008), MOTHUR (Schloss et al., 2009), LIBSHUFF (Cardoso and Coutinho, 2012b), MrBayes 3.2 (Ronquist et al., 2012), TreeOTU (Tree-Operational Taxonomic Unit) (Wu et al., 2013), DAMBE5 (Xia, 2013), PhyloTreePruner (Kocot et al., 2013), BlastGraph (Ye et al., 2013), LSCluster (Large-Scale CLUSTERing) (Husi et al., 2013) and many others have been designed to determine the systematic distances between two or more given sequence libraries. This software uses statistic calculation processes, such as principal coordinate analysis and clustering, to calculate the evolutionary distances between the members of two communities and to draw their phylogenetic trees (Wooley and Ye, 2010). The UniFrac programme determines the phylogenetic distances in a multiple database in order to compare the differences between microbial communities of two or more habitats. Furthermore, this approach is able to show the effects of different factors, eg, geography, chemistry and pollution using, on a given community (Lozupone and Knight, 2005). DOTUR is a software used to quantify the

sensitivity of the metagenomic approach for determination of the microbial diversity in a habitat, and to show what is required to achieve full sequence coverage of the diversity of a community (Rasheed et al., 2013, Rasheed et al., 2012). MOTHUR is a bioinformatic package that collects different software, including DOTUR and UniFrac, as a combined software to simplify analysis of microbial communities (Schloss et al., 2009). Although most of these softwares are designed to determine the similarities between 16SrRNAs, this method is not perfect due to horizontal gene transfer phenomenon and variation in the number of 16sRNA genes among microorganisms. To address this problem, TreeOTU has been supported with an algorithm which detects the similarities between protein-coding gene families as an evolutionary marker (Wu et al., 2013). The newest softwares, such as DAMBE5, involve many other biological features including rRNA, tRNA, exons, introns, consequence sequences of genes (like promoter, operator, terminator and first codons of genes), secondary structures of RNAs and many other properties in their algorithms to resolve the existing limitations in phylogeny of microorganisms and to design highly accurate phylogenetic softwares (Xia, 2013). Although these methods are quite accurate and specific, they are not useful for all metagenomic studies due to the possible uniqueness of the genomic sequence of unknown microorganisms, and also to their complexity and high cost (Mohammed et al., 2011, Husi et al., 2013).

Composition-based methods such as TETRA (Teeling et al., 2004), Phylopythia (Rosen et al., 2011), signature sorting based assignment procedure (SSBA) (Rachamalla et al., 2012), Metagenome Composition Vector (MetaCV) (Liu et al., 2013), however, use intrinsic compositional structures of the given sample including oligonucleotide frequencies of the genome, GC content and codon usage preferences to find the taxonomic category of the sample (Liu et al., 2013). Although composition-based methods are faster and lower cost than the alignment based method, they are less accurate (Diaz et al., 2009, Brady and Salzberg, 2009). Furthermore, sequences shorter than 1 kbp (in conventional sequencing methods) or shorter than 100 bp (in next generation sequencing approaches) have insufficient signals for taxonomic investigations (Brady and Salzberg, 2009). In order to improve the approaches in terms of accuracy and time consumption, many hybrid methods such as SPHINX (Haque et al., 2009, Mohammed et al., 2011), MetaSAMS (Zakrzewski et al., 2012) and phiSpy (Liu et al., 2013) have been designed in which the data are analysed by compositional features before investigation based on BLAST.

It is important to note that, in addition to phylogenetic analysis, many metagenomic investigators aim to detect the presence of specific abilities of particular microorganisms. The gene prediction softwares are classified into three groups: 1) homology-based programs, such as PipMaker (Schwartz et al., 2000), SGP-1 (Wiehe et al., 2001), AGenDA (Taher et al., 2003), GENEWISE (Birney et al., 2004) and BLANNOTATOR (Kankainen et al., 2012) use gene sequence databases such as Smith-Waterman algorithms and BLAST for prediction of the closest homologous sequence. 2) splicing site prediction softwares, such as NetGene2 (Hebsgaard et al., 1996), Genesplicer (Pertea et al., 2001), ESEfinder (Cartegni et al., 2003), HSF (Human Splicing Finder) (Desmet et al., 2009), work based on maximal dependence decomposition (MDD) and are specialized for eukaryotic organisms. 3) the Ab initio softwares, such as GENIE (Reese et al., 1997), Gene-Mark (Lukashin and Borodovsky, 1998), GeneBuilder (Milanesi et al., 1999), GLIMMER (Delcher et al., 1999, Kelley et al., 2012), GENSCAN (Cawley, 2000, de Marvao et al., 2013), ASPic-GeneID (Alioto et al., 2013) and MetaGeneTack (Tang et al., 2013) use statistic models for predicting the genes based on finding consensus sequences in the Open Reading Frames (ORFs) such as promoters, exons and introns (Yok and Rosen, 2011). The Newest structure based programs, such as MESSA (MEta-Server for protein Sequence Analysis) (Cong and Grishin, 2012), FFPred 2.0 (Minneci et al., 2013) exploit many other molecular features such as secondary structure, transmembrane helices, signal peptides, 3D structure and structurally disordered regions to facilitate gene and protein analysis. FFPred 2.0, also, is able to predict ontology of genes, their role in metabolism, signalling and regulatory events (Minneci et al., 2013).

Despite the sequence based software, which are limited to known genes (She et al., 2009, Yok and Rosen, 2011), other gene prediction software, such as GENSCAN, GENIE, GLIMMER and Gene-Mark.hmm, are able to predict the existence of genes based on finding the Open Reading Frames (ORFs); each ORF has an initial (AUG) and terminal codon (UAA, UAG and UGA). A technical problem of such systems is that since most of the reads are quite short (less than 500 bp) and may contain only partial segments of genes, most open reading frames are ignored (Hoff et al., 2009). This happens because these systems are based on statistical sequence models and rely on genomes already analysed, and most reads lack useful informative sequences for statistical analysis (Hoff et al., 2008). This problem is resolved in a two-step process called MetaGene, which first recognises ORFs and determine their scores, and then classifies the gene candidates based on the GC continents, ORF length and the distance between neighbouring ORFs (Noguchi et al., 2006, Hoff et al., 2008).

In order to facilitate metagenomic studies for operators, new versions of software are able to determine both phylogeny and gene contents of a given sample. MEGAN 5 (MEta Genome ANalyser) (Tamura et al., 2011), for instance, is a powerful approach for studying the taxonomic content of multiple datasets and for determining functional analysis of the data. Furthermore, MG-RAST (Metagenomics–RAST) is designed as a web-based approach (http://metagenomics.anl.gov/) which in addition to phylogenetic analysis is able to identify the metabolic pathways used for the degradation of the substrate of interest by analysing operons or reorganising the data in order to find the genes involved in particular metabolic pathways (Glass et al., 2010, Wilke et al., 2013).

2.3.4 Applications of metagenomics in ecological studies

The traditional culture based discovery approaches for finding new genes of interest or their products are restricted to those microorganisms that can be grown on suitable media. Metagenomic analysis using high throughout sequencing approaches and powerful bioinformatic softwares, however, enables identification of both the microbial diversity and genes of interest without this restriction. Table 2.3.4 provides a comparison of traditional culture-based and molecular-based technologies and metagenomic studies for the investigation of microbial diversity and gene discovery.

Microbial ecology aims to link the microbial diversity in a given environment to the natural processes performed in that environment. In this case, since metagenomic studies provide information on both the microbial diversity and the genetic capacity of the environment, it is a powerful tool for analysis of the physiological abilities of microorganisms in a given environment. Highly conserved genes or anchors within a microbial genome, especially the V3 + V4 regions of the 16S rRNA gene which are used for the identification of microbial diversity in a community without requiring information of the full genome sequence (Klenk and Göker, 2010). In the case of prokaryotes, rDNA sequence with more than 97% similarity are clustered as a species or an Operational Taxonomic Unit (OUT) (Chang et al., 2008, Turnbaugh et al., 2009). Due to the limitations of traditional methods and to extreme conditions in some cases, outcomes of metagenomic studies are more reliable (Chang et al., 2008, Turnbaugh et al., 2009). For instance, extreme changes in physical and chemical environmental properties, including pH, temperature, pressure, concentrations of salt or heavy metals, and radiation create a variety of environmental pressures such that only specially adapted microorganisms

(extremophiles) are able to colonize such environments (Inskeep et al., 2010, Liu et al., 2011e). Low microbial diversity and the simplicity of the relationships between the microorganisms in these extreme conditions enhance the ability of metagenomic approaches to identify the constituent members of the community and to detect their ecological roles in extreme environments (Lewin et al., 2013). For instance, metagenomic studies on microbial biofilms of hydrogen sulfide-rich caves (pH 1) indicated that *Acidithiobacillus thiooxidans* is the predominant microorganism, which uses sulphide-quinone reductase to extract electrons from hydrogen sulfides and transfer these electrons via sox pathways to a RuBisCO system (ribulose-1,5-bisphosphate carboxylase/oxygenase) to fix CO_2 (Jones et al., 2011).

Iron oxidation in pyrite by microorganisms in mines leads to acidified mine drainage into the surrounding lands (Edwards et al., 2000). The simplicity of the microbial community in acidic mine drainage (AMD) water and the considerable differences in GC% of the organisms simplified the genomic reconstruction process in this case. A metagenomic study on microbial biofilms of AMD (pH 0.83) could detect members of *Sulfobacillus* sp., *Leptospirillum,, Ferroplasma* sp., G-plasma, A-plasma as well as some protozoa containing *Rickettsial* endosymbionts (Baker et al., 2003). These studies were also able to investigate the ecological roles of the members of the AMD microbial community and the different genes responsible for CO_2 fixation, microaerophilic survival, sugar, amino acid transporters and biofilm formation were identified (GW et al., 2004).

Purposes		Culture based approaches	PCR based molecular approaches	Metagenomic studies
Microbial diversity	Advantages	Gold standard technology due to isolation of microorganisms	1) Accurate (but not 100%) 2) fast	1) enable to sequence a mixture of microorganisms 2) High-throughout approach 3) Low errors 3) Employment of other consensus marker genes in addition to 16SrRNA for identification of microorganisms
	Disadvantages	1) Time consuming 2) preparation of several media and chemical reagents for detection of the organisms 3) is able to support only 1% of microorganisms	1) availability of a pure culture or isolate 2- Sequencing errors as a result of Taq DNA polymerase activity, chimera formation and primer biases	Expensive
Enzyme discovery	Advantages	----------	1) Detection and isolation of whole gene or operon of interest	1) detection of almost whole gene capacity 2) designing a specific primer for a new gene based on the sequence obtained from high-throughout sequencing
	Disadvantages	Only enables to isolate enzyme producing microorganisms not the genes	1) Unable to detect different genes in one experiment 2) It is a blind tool because of employment of a known universal primer specific for gene(s) of interest 3) missing most of the clone harbouring colonies because of limitation in selection of colonies	1) Production of short length reads which can increase error rates 2) Expensive

Table 2.3.4. Comparison of traditional tools and metagenomic-based approaches in identification of microbial population in a community and detection of gene(s) of interests.

Furthermore, metagenomic studies have shown that although the microbial composition and species abundance of different high-thermal environments (>45 ^{o}C) are not constant, species from some bacterial taxa (e.g. *Deinococcus-Thermus*, *Thermotoga* and some proteobacteria) and archaeal taxa (e.g. *Thermococcus*, *Methanococcus*, *Desulforudis* and *Thermoprotei*) are found in most of these thermal environments (Kotlar et al., 2011, Nelson et al., 2011, Xie et

al., 2010). Overall, such extreme habitats are not rich in terms of species diversity and only a few taxa are dominant in these environments. These studies also showed that thermophilic microorganisms use different strategies, such as configurationally heat stable cellular proteins (with a higher contents of β-plated sheets and more covalently linked cross bonds), heat-shock proteins (HSPs), saturated phospholipids (bacteria) or phosphoesters (archaea) in cell membranes and high CG% genomic material for adaptation to the extreme conditions (Szilágyi and Závodszky, 2000, Dalhus et al., 2002). High expression of radA and DNA damage repair systems in these inhabitants indicate that they need to monitor and repair DNA damage (Seitz et al., 1998, Trivedi et al., 2005). Furthermore, thermophilic organisms often express a reverse DNA gyrase that stabilises the chromosome around nicks via introduction of positive supercoils to DNA (in comparison to most gyrases that introduce negative supercoils in mesophiles) (Trivedi et al., 2005).

Low temperature environments (below +5 ^{o}C) cover the largest portion of the Earth's biosphere (Bartlett, 2008). Since these areas are often restricted in terms of nutrients, the psychrophile or *cryophile* communities reveal an oligotrophic or autotrophic metabolism in response to the limited nutrient sources (Bartlett, 2008, Tehei et al., 2003). Furthermore, these microorganisms use different types of psychrophilic adaptation system, including production of cryoprotectants (anti-crystallizer), increasing unsaturated fatty acids in the cell membrane to maintain the cell membrane fluidity, production of exopolysaccharides, expression of cold shock proteins and the elimination of reactive oxygen from the cell environments (which can be high as a result of increases in O_2 solubility), as adaptations to such low temperature systems (Bartlett, 2008, Yamada et al., 2000, Chattopadhyay, 2006). Based on metagenomic studies in different cold regions, proteobacteria and cyanobacteria dominate habitats in the ice shelves of Antarctica, while green sulphur bacteria (*Chlorobium* sp.) and sulphate reducing bacteria (SRB) are the most frequently detected organisms in Alaska and the Canadian High Arctic (Varin et al., 2012, Rankin et al., 1999). These two groups of bacteria live in a close relationship together; SRBs reduce sulphate to H_2S that is absorbed by chlorobi species to supply their reductive energy and electrons. The sulphate released by chlorobi is utilised by the SRBs continuing the cycle (Rankin et al., 1999, Ng et al., 2010).

In addition to the efficiency of metagenomic studies in the analysis of microbial diversity, these approaches are able to determine the physiology and metabolism of microorganisms in an environment. The physiological features and biochemistry of most microorganisms have not

been discovered since most of these organisms are not culturable on any laboratory media (Vieites et al., 2009). For instance, although eight groups of the *Acidobacterium* branch have been identified and its members are distributed worldwide, only three of the groups are culturable (Sait et al., 2002, Liles et al., 2003). However, based on information on functional traits obtained from metagenomic approaches and using enrichment of samples with certain substrates, especially in combination with stable isotope probing (SIP) or by providing selective media with special chemical or physical demands, it is possible to design a specific media for microorganism of interest and thus to improve the scientific approaches for the investigation of such microorganisms (Joshi et al., 2014, von Bergen et al., 2013).

Furthermore, metagenomic studies are also able to investigate the evolutionary processes in a sympatric population, where the same genes can be shared among microorganisms from different phyla, and therefore have helped to reveal the existing biases in the classification of microorganisms. Based on studies performed on the genomes of Prochlorococcus populations of the Atlantic and Pacific oceans, for instance, the highest level of genetic variation arises in genomic islands, partly due to gene transfection by viruses (Hugenholtz and Tyson, 2008, Polz et al., 2013). In another case, despite the initial identification of Rhodopsin as a phototrophic molecule in Archaea, an uncultured γ-Proteobacterium harbouring the same molecule, referred to as bacteriorhodopsin, was found (Sabehi et al., 2003, Béja et al., 2000). Similar studies on archaea and their relationships with different organisms in a variety of terrestrial and aquatic environments have elucidated their physiological and ecological roles in their environments. In one of the studies, for instance, it has been shown that high frequency genetic recombination in archaeal communities present in acid mine drainage is the main evolutionary force creating diversity in these populations (Denef et al., 2010, Tyson et al., 2004).

2.3.5 Applications of metagenomics in clinical studies

Metagenomics is also a powerful approach for the direct detection of microorganisms in clinical samples and for tracking infectious microorganisms in different environments. Many reports have been already published emphasizing the ability and accuracy of metagenomic studies in the identification of infectious diseases (Nakamura et al., 2008, Nakamura et al., 2009). Theoretically, the metagenomic approaches are reliable ways to detect the causative agents of several types of culturable and non-culturable infectious disease (Nakamura et al., 2008). The future perspectives in clinical microbiology highlight the role of metagenomic approaches in the generation of a massive amount of genetic information from known/unknown

microorganisms and links between this information and several human diseases (Mokili et al., 2012). For instance, a normal viral metagenome data obtained from an environment like a human being contains 60-99% unknown data which, regardless of the technical problems for the production of these sequences, suggests at least a small group of these sequences can be associated with unclassified viruses (Mokili et al., 2012).

In addition to the direct identification of microorganisms in clinical samples, the metagenomic approaches facilitate microbial tracking in clinical or environmental samples relying on similarities and dissimilarities between different molecular markers among the microorganisms found in clinical samples and the environment. Employment of metagenomic approaches to find the origin of an Outbreak of Hemolytic–Uremic Syndrome in Germany, for instance, showed that the causative *E. coli* O104:H4 strain emerged as a result of gene recombination between a Shiga-toxin–encoding phage and an enteroaggregative *E. coli* strain (Rasko et al., 2011). Using a metagenomic study on methicillin-resistant *Staphylococcus aureus* (MRSA) strains isolated from patients, it was possible to distinguish between the strains and therefore, to investigate a MRSA outbreak in a neonatal intensive care unit (Köser et al., 2012). Furthermore, based on the reassembling of the resistome of genes coding for antibiotic resistance in these strains, they could determine an antibiotic resistance pattern, which was more informative than the normal antibiogram tests (Köser et al., 2012). Since environmental sources are considered as the main reservoirs of different microorganisms and therefore can provide a substantial potential for infecting target organisms (human beings, animals and plants) exposure, identification of these contaminated sources can be helpful in providing protection in that area. Many viruses with the ability to resist harsh environmental conditions, especially Polioviruses, Hepatitis A and E viruses, Adenoviruses and Reoviruses, are able to survive for a long time in these reservoirs and therefore can create an outbreak in the case of exposure to humans. In these cases, a complete sequencing technique based on metagenomic study is a promising assay for detection of such contaminated sources (Julian and Schwab, 2012). Fecal indicator microorganisms are usually used to test for any water or food contamination originating from human or animal faeces. In addition to the identification of the microorganisms in a given sample, the metagenomic approach is able to differentiate the environmental microorganisms (viruses and bacteria) from the stool-originated microorganism (Stachler and Bibby, 2014). In a 16S rRNA gene library-based examination for determining the origin of faecal bacteria in several French sewages, for instance, this technique could show that all isolated *Clostridiaceae* and *Bifidobacterium* originated from human stools while

Bacterioidales had both faecal and non-faecal origins (Wéry et al., 2010). In another metagenomic study a highly sensitive faecal indicator microorganism, CrAssphage was introduced as a suitable candidate since it was only found in the human gut and bat guano and could be tracked in the environmental samples (Stachler and Bibby, 2014).

Despite the risks emerging from the antibiotic resistance ability of microorganisms, there is no efficient knowledge about the origin, distribution and diversity of the resistance genes, especially among unculturable microorganisms (Garmendia et al., 2012). The application of non-culture based approaches opens a new window for investigations on the presence of both antibiotic production and antibiotic resistance in different ecosystems both in order to find new sources of antibiotics and to predict the emergence of antibiotic resistance in microorganisms (Garmendia et al., 2012). In a study on remote Alaska soil, for instance, the presence of divergent b-lactamases was proven even in the absence of selective pressure (Allen et al., 2009). The same studies of sewage showed that many of the heterotrophic bacteria in rivers and their inhabitants, including fishes, showed multi-antibiotic resistance to many existing antibiotics (Garcia-Armisen et al., 2011, Verner-Jeffreys et al., 2009).

With the advent of high-throughput sequencing it is also possible to investigate the whole microbial population associated with the human body (Hamady and Knight, 2009). Several institutes worldwide participate in the Human Microbiome Projects (HMPs) in which metagenomic studies and many other techniques are used to identify these populations and their roles in human health and diseases (Consortium, 2012). The main questions investigated by these institutes is the possible presence of a core human microbiome of microbial genes found in all human beings and changes in their proportion in disease conditions (Consortium, 2012). Metagenomic studies on the microbial populations of the human gut, for instance, showed non-random distribution patterns in different parts of the gastrointestinal tract, and even this is variable in different individuals and at different ages (Ley et al., 2006, Abbasian and Saberbaghi, 2013b). Globally, the symbiotic or commensal relationships between host and microorganisms, especially *Bifidobacterium* and *Lactobacillus*, begins from the first days of birth as a result of passing through the maternal birth canal and of breastfeeding, and is necessary in order to maintain good health (Hooper and Gordon, 2001). Later, when infants start eating food other than milk, *Escherichia coli, Bacteroides, Bifidobacterium, Streptococcus, Enterococcus, Clostridium* and *Actinomyces* become the dominant intestinal organism (Favier et al., 2003). This variability in the intestinal microbiome can probably affect

a variety of physiological activities in the intestinal tract and extra-intestinal compartments of the human body. Many researches have been performed on the influences of intestinal microbiome on development of brain cells, personal behaviour and overall psychological reactions (Heijtz et al., 2011, Cryan and Dinan, 2012). The intestinal microorganisms influence the brain through changes in the function of the hypothalamic-pituitary-adrenal axis and the cortex, effects on the neurochemical secretions, modulation neuronal transmitters such as gamma-aminobutyric acid, acetylcholine, histamines, serotonin and melatonin) (Wikoff et al., 2009, Uribe et al., 1994). There are direct correlations seen between the intestinal microbial diversity, especially in the number of aerobic bacteria, and several psychiatric disorders. For instance, appearance of autism and major depressive disorder (MDD) are associated with increases in the abundance of clostridia, specially *Clostridium bolteae*, (Finegold, 2008) and different aerobic bacteria (Logan and Katzman, 2005), respectively. However, it has not yet clarified whether changes in the biodiversity of microorganisms in a physiological condition is a side effect of these disorders or are the real causative factors of the disorders.

2.3.6 Application of metagenomic approach in biotechnology

The capability of microorganisms to survive and grow in different environments is related to microbial physiology and their metabolic capacity (Schimel et al., 2007). Since metagenomic tools provide the ability to investigate the genetic capacity of these microorganisms, it is possible to identify genes of biotechnological interest and transfer them to a fast growing microorganism, such as *E. coli*, and therefore, to employ them in different industries (Faheem et al., 2013). The outcomes of metagenomic studies can be used for several biotechnological purposes, such as discovering new antibiotics or new enzymes and for the remediation of natural and artificial pollutants. Most industrial and municipal anthropogenic activities depend on the utilization of thousands of types of natural and artificial chemicals and massive quantities of energy, which ultimately lead to the production of large amounts of discharges into the environment (Mani and Kumar, 2014, Gołębiewski et al., 2014). This situation highlights the serious and urgent actions needed to clean up the resulting contaminated sites using different mechanical, chemical or biological approaches. Among them, bioremediation using suitable microorganisms is a promising technique for the removal of the contaminants while leaving the lowest side effects in the area (Mani and Kumar, 2014). However, since the functional potential of microorganisms have not been completely understood, culture independent technologies, such as metagenomics, are very helpful to find proper microbial strains or genes involved in the removal of contaminants (Haiser and Turnbaugh, 2013).

Hydrocarbon pollution, for instance, is one of the common adverse outcomes of industrialization coming from a range of sources, including oil spills (Margesin and Schinner, 2001a), and affects many different organisms in both aquatic and soil environments. Most oil spilling events occur accidentally in petroleum refineries, storage facilities and oil terminals (Chang and Lin, 2006). A large diversity of microorganisms has been identified worldwide that are able to remediate these contaminants by biodegradation. Metagenomics is potentially applicable for research on microbial succession in oil contaminated microcosms and for the discovery of novel types of aliphatic and aromatic hydrocarbon biodegradating enzymes (Cardoso and Coutinho, 2012a). These studies indicated that several groups of prokaryotic (*Helia sp.*) and eukaryotic microorganisms are susceptible to oil contamination, being lost from the microbial community. In another study (Gutierrez, 2011), found that although metazoans and fungi were the dominant microorganisms in the soil based-microcosm studied, oil contamination decreased their abundance and they were replaced by *Stramenopiles* as the dominant organisms after hydrocarbon contamination. As a result, the equilibrium between metazoan and fungi on the one hand and *Stramenopiles* on the other hand can be used as an oil spillage indicator that reflects the level of oil toxicity on the studied ecosystem (Cardoso and Coutinho, 2012b). Based on information from conventional and newly emerged technologies, several microorganisms, such as *Pseudomonas aeruginosa*, *Stenotrophomonas* sp. and *Ochromonas* sp. (Ueno et al., 2007) are able to degrade oil components. These studies also confirmed the role of monooxygenase and dioxygenase genes as the initiator enzymes for the catabolism of polycyclic aromatic hydrocarbons (PAH). These genes catalyse a crucial step in PAH remediation pathways through the insertion of one or two oxygen atoms into benzene rings (Marcos and colleagues (Ono et al., 2007).

Screening of DNA libraries based on sequence homology, function-based screening and gene expression influenced by substrate dependent induction, is a powerful approach for discovery of many of the genes for the production of metabolites (such as anti-cancer drugs and antibiotics) and enzymes of interest (von Bergen et al., 2013, Seifert et al., 2012). Since the production of some secondary metabolites provides an advantage in environmental competition to particular microorganisms, it is reasonable that such metabolites are to be found in many natural terrestrial and aquatic environments, especially in those environments with a complex microbial community, such as soil, animal's gastrointestinal tracts, aerobic and anaerobic sludge in wastewater treatment systems. For instance, based on sequence homology-based

metagenomic studies, an anticancer product (pederin) originally identified in *Paederus fuscipes* (Joshi et al., 2014), was detected in a symbiotic *Pseudomonad* (Xiong et al., 2013). Metagenomic approaches were also able to discover genes for the synthesis of many secondary metabolites, including antibiotics (Xiong, 2013) and polyketides (Trindade-Silva et al., 2013). Discovery of several novel anti-mycobacterial or anti-fungal antibiotics, such as terragine and turbomycin A and B is an excellent example in which metagenomic approaches led to the discovery of novel antibiotics (King et al., 2009). In addition to the discovery of novel antibiotics, metagenomic studies could help biotechnologists to improve the antimicrobial activity of existing antibiotics. Peptide synthetases and polyketide synthases (PKSs), for instance, are two critical enzymes involved in production of complex antibiotics (Woodhouse et al., 2013, Della Sala et al., 2013). PKSs consist of a divergent repeating domain, responsible for production of antibiotics with different chemical structures, flanked by extremely conserved regions (Della Sala et al., 2013). In practice, these conserved regions are first used as a template for designing the probes that detect PKS genes in metagenomic studies, and then different PKS domains obtained from bacterial species are combined together in order to generate new effective antibiotics (Williams, 2013). Finally, the genes involved in the biosynthesis of different vitamins, such as Riboflavin (vitamin B2), vitamin B12, other Vitamin B groups and vitamin K are another target for metagenomic studies (LeBlanc et al., 2013). A metagenomic study on soil samples, for instance, was successful in detecting the genes responsible for encoding 2,5-diketo-D-gluconic acid reductases, which are responsible for the conversion of glucose to of vitamin C ((ascorbic acid) (Eschenfeldt et al., 2001).

Furthermore, such metagenomic studies using enriched or selected media facilitate studies on the metabolic activities of a community and these have led to discovery of many new varieties of enzymes of interest in industries and medicine such as cellulases (Lepage et al., 2013), lignocellulose metabolising enzymes (Nimchua et al., 2012), pectinolytic lyases (Wang et al., 2014a), acidic thermostable α-amylases (Sharma et al., 2010, Jabbour et al., 2013), b-glucosidase (Wang et al., 2012a), glycosyl hydrolase (Sathya et al., 2014), xylan hydrolysing enzymes (Yang et al., 2011), lipolytic enzymes (Zhang et al., 2013a), flavonoid modifying enzymes (Rabausch et al., 2013), nitrilases (Gong et al., 2013) and so on. A metagenomic study on the hindgut of a wood-feeding higher termite, as a successful wood-degrading organism, revealed that phyla Treponema and Fibrobacteres are dominant in the microbial diversity in that environment, and they contribute the majority of cellulolytic genes in the metagenomic DNA (68% and 13%, respectively) (Warnecke et al., 2007). In addition, a total of 34 genes of

unknown function were found in this experiment, which presumably are involved in binding to the substrates and some other functions responsible for lignocellulose conversion (Warnecke et al., 2007). In comparison, similar metagenomic study on another animal's gut microbiome found a diverse variety of genes encoding cellulolytic enzymes (Brulc et al., 2009, Zhu et al., 2011). In addition to the sequence diversity of enzymes identified in the metagenomic studies, the catalytic rate and specificity of many of these genes are different from known genes in the gene databases. Therefore, finding a dozen variation of a special enzyme with the ability to display different optimal pH, temperature and ionic conditions can be useful for the application of the enzymatic reaction in a large range of physical conditions (Steele and Streit, 2005). In a metagenomic study on soil samples, for instance, an amylase encoded by an uncultured bacterium was cloned into *E.coli* which retained 90% of its maximum activity at a temperature range of from 10-30 ^{o}C (Sharma et al., 2010). As well, a novel alkaline lipase with no amino acid similarity was found in a metagenomic investigation which was adapted to in low temperatures (Kim et al., 2009). Overall, All these examples emphasize the efficiency and usefulness of metagenomics in providing a more realistic understanding of the activity and diversity of microorganisms in any particular environment and the ability of this approach to provide information on the diversity of new microorganisms and gene(s) of interest that may be useful in different industries.

2.3.7 Conclusion

Metagenomics is a new promising approach for investigation of biodiversity and is applicable to gene pool of an environment. One advantage of metagenomics is that direct sequencing of the DNA does not need any prior knowledge of the gene sequence, and it is theoretically possible to access the sequence of genes and thereby determine the function of these genes. However, since metagenomics is highly dependent on the nucleic acid sequences of given sample, improvement of the sequencing method is critical for more accurate analysis. Research in this area should aim to decrease the cost of a sequencing reaction and the time consumed for each reaction. Furthermore, since the sequencing reaction release massive amounts of data, employment of bioinformatic programs is necessary for their analysis. Although there are many softwares applicable for analysis of metagenomic data, analysis speed, qualification of analysis process, the ability to link to different gene databases and user friendly environment are the most important options for superiority of a software.

The metagenomic approach is a powerful tool to investigate the microbial diversity in an environment, and to determine the ecological roles of the microbes present in a habitat. The outcomes of these investigations can also be used in biotechnology to find new enzymes and antibiotics and to resolve several environmental issues using microorganisms, such as global warming and cleaning up of the environment. Environmental pollution caused by the recalcitrant substances that are resistant to biodegradation and also contamination of extreme environments by pollutants might be resolved by finding new microorganisms or enzymes. It is worthwhile to mention that emerging technologies are usually associated with contamination with new products, such as artificial nanoparticles, for which there is no known degrading system in nature. In these cases, microorganisms may need a long period of time before they adapt by evolving new pathways for the degradation of such compounds. However, metagenomics may bypass that limitation and can facilitate the identification or evolution of effective microorganisms or enzymes for degradation, or at least for neutralization of the detrimental effect(s) of the emerging pollutants.

Section 3: Experiments

3.1 A pyrosequencing-based analysis of microbial diversity governed by ecological conditions in the Winogradsky column

Published in: World Journal of Microbiology and Biotechnology; 31(7): 1115-1126.

3.1.1 Abstract:

The Winogradsky column is used as a microcosm to mimic both the microbial diversity and the ecological relationships between the organisms in lake sediments. In this study, a pyrosequencing approach was used to obtain a more complete list of the microbial organisms present in such columns and their ratios in different layers of this microcosm. Overall, 27 different phyla in these columns were detected in these columns, most (20 phyla) belonged to bacteria. Based on this study, *Proteobacteria* (mostly *Sphingomonadales*), Cyanobacteria (mostly *Oscillatoriales*) and Bacteroidetes (mostly *Flavobacteriales*) were the dominant microorganisms in the water, middle, and bottom layers of this column, respectively. Although the majority of organism in the water layer were photoautotrophic organisms, the ratio of the phototrophic organisms decreased in the lower layers, replaced by chemoheterotrophic bacteria. Furthermore, the proportion of aerobic chemoheterotrophic bacteria was greater in the higher layers of the column in comparison to the bottom. The green and purple sulfur phototrophic bacteria inhabited the bottom and middle of these columns, with none of them found in the water layer. Although the sulfur oxidizing bacteria were the dominant chemolithotrophic bacteria in the water layer, their ratio decreases in lower layers, being replaced with nitrogen oxidizing bacteria in the middle and bottom layers. Overall, the microbial population of these layers changes from a phototrophic and aerobic chemoheterotrophic organisms in the water layer to a mostly anaerobic chemoheterotrophic population of bacteria in the bottom layers.

3.1.2 Introduction:

Black mud in a lake or sea is an almost anaerobic environment containing a variety of organic and inorganic sulfur compounds as well as iron (Tanner et al., 2000). Anaerobic microbial respiration in these conditions leads to the production and precipitation of insoluble black FeS (Enning et al., 2012) and sulfur compounds are also oxidized as a result of the activities of a

variety of chemolithotrophic and photoautotrophic microorganisms (Basen et al., 2011, Dutta et al., 2014). Upper layers of the black sediment and the water layer can be inhabited by a variety of photoautotrophic and chemoheterotrophic microorganisms depending on the availability of O_2 and light at various depths (Dutta et al., 2014, Tanner et al., 2000). At the expense of CO_2 reduction , the activities of the photoautotrophic organisms provide the organic substrate required for chemoheterotrophic microorganisms in lower layers (Kirchman, 2010). One group of these phototrophic bacteria, referred to as photolithotrophic bacteria, is able to reduce CO_2 through the oxidation of H_2S (Kirchman, 2010, BELL, 2013). On the other hand, the respiratory activity of anaerobic sulphate reducing bacteria on organic compounds returns the CO_2 and H_2S required for photolithotrophic bacteria (Pereira et al., 2011). The relationships between different microorganisms in an environment lead to the formation of sulfur and carbon cycles in that environment (Paul, 2014). Such microbial cycling of elements normally occurs in different soil and sediment environments (Paul, 2014). However, the composition and activity of these microbial communities are affected by different physical and chemical changes in the environment, such as light intensity, temperature, pH, availability of water and nutrition, availability of molecular oxygen and presence of toxins (Ganzert et al., 2011, Rousk et al., 2010, Zhong et al., 2010). Since *in situ* studies of these complex environments are very difficult, microcosms are the best way to mimic these environments.

The Winogradsky column is a suitable column for such studies designed by Sergei Winogradsky in the late 1800's to study the ecological conditions of sediment microorganisms (Dworkin and Gutnick, 2012). This column is made of a transparent chamber filled with sediment, water and air layers, which mimics the geobiological conditions of a shallow lake (Teramae et al., 2012, Tanner et al., 2000). The bottom sediment layer is enriched by cellulose and sulphate salts to increase the activity of sulphate reducing bacteria (SRB) (Teramae et al., 2012, Tanner et al., 2000). Energy and carbon sources required for growth of microorganisms are provided by photoautotrophic activities or the organic molecules already present in the sediments (Moshynets et al., 2013). The light irradiation from above and the concentration of sulphide from below in addition to the concentration of free oxygen (O_2) are the three main differential factors which limit microbial activities in such a column (Hounshell, 2012, Temkar, 2011). Based on these limitations, a gradient of microbial biodiversity occurs within the different layers of the microcosm (Hounshell, 2012, Temkar, 2011).

Although a list of microbial species present in the different layers of this kind of column have been published in several papers (Tanner et al., 2000, Teramae et al., 2012, Rundell et al., 2014), most of these analyses used conventional approaches such as microbial culture and clone library based 16S rRNA analysis. Since culture-based methods are able to support the growth of only 1% of a microbial population (Vieites et al., 2009, Dionisi et al., 2012) and the construction and sequencing of genome libraries limits the analysis to a small portion of the bacterial clones, these methods are not suitable for the complete analysis of the microbial diversity in a complex environment (Tian et al., 2010). However, since the column is made with soil and sediment samples, a very complex microbial diversity is expected in the different layers of the Winogradsky column. Rundell and colleagues (Rundell et al., 2014) performed a perspective Illumina's MiSeq sequencing-based study on the microbial diversity of this column and the effects of some physical and chemical factors on this diversity. They confirmed the presence of highly diverse microbial communities at different depths in the Winogradsky columns and showed the relative impacts of the source of the sediment, depth and supplemental cellulose source on these communities. Furthermore, they showed that the microorganisms involved in recycling sulphate, including both sulphate reducing bacteria (SRBs) and *Desulfobacteraceae* are dominant in these environments. However, further studies are required to determine the overall physiological features of microbial community and the ecological relationships governing the relationships between microorganisms at different depths in the column. In this study we tried to answer these questions using a high throughout sequencing approach.

3.1.3 Material and method

3.1.3.1 Preparation of the Winogradsky column

Two Winogradsky columns (9 cm in diameter and 25 cm high) were prepared from a pond located at Mawson lakes (Adelaide, Australia). The columns were created following the methods used in previous studies (Anderson and Hairston, 1999, Loss et al., 2013). Following collection of the sediments from the pond (under roughly 15–30 cm of water) in July 2014, the sediments were sieved to remove rubble, roots and macroscopic animals.

To set up the bottom of the Winogradsky columns, a portion of the sediment was mixed with 0.25% W/W $CaCO_3$, 0.50% W/W Na_2SO_4 and around 20g of shredded newspaper. One third

of the columns was filled with this combination while the second third was filled with sieved sediments only. Finally, one sixth portion of these column was filled with the lake water. To prevent any light irradiation from the sides, two thirds of the columns from the bottom was covered with aluminium foil. The columns were incubated in conditions where the top of the columns were constantly exposed to normal lamp light (400–700 nm wavelength). The columns were incubated for 30 days at a temperature of 25 °C.

3.1.3.2 DNA extraction and Pyrosequencing process

Samples from three different layers (the bottom, the middle and the water layers) were obtained by passing syringe needles (nominal inner diameter 3.81 mm**)** through the plastic walls of the columns. To improve the sampling accuracy, three samples were taken from different parts of each layer of the columns and then the samples obtained from each layer were mixed to give a homogenous sample for each layer. DNA extractions were performed using the power soil DNA kit (MO BIO), according to manufacturer's instructions in which 0.25 g sediment samples were added to the powerbead tubes, and after a gentle vortex were treated with different solutions. First, 60 µl of solution C1 was added to each tube and the tubes were vortexed vigorously for 10 minutes. These tubes were next centrifuged at 10,000 x g for 30 S and the supernatant (400-500 µl) were transferred to a clean 2-ml tube. Following addition of 250 µl of Solution C2 and a short time (5 S) vortex, the tubes were incubated for 5 minutes at 4 °C and were centrifuged at 10,000 x g for 1 minute. Up to 600 µl of the supernatants were transferred to new 2-ml tubes and the tubes were treated with 200 µl of Solution C3. Following a brief vortex and an incubation period at 4 °C for 5 minutes, the tubes were centrifuged at 10,000 x g for 1 minute. Once again, up to 750 µl of supernatants were transferred into another clean 2-ml tubes and the contents were mixed with 1200 µl of solution C4. Approximately 675 µl of these mixtures were loaded onto a spin filter and following a centrifugation at 10,000 x g for 1 minute, the flow were discarder through. This loading and centrifugation step were repeated for three times to separate all DNAs from the solution. A total of 500 µl of solution C5 was added to each spin filters, and after a centrifugation at 10,000 × g for 30 seconds, the flows were discarded; these centrifugation steps were repeated twice to remove all flows from the filters. The spin filters were placed in a new 2-ml tube and the DNAs were eluted from the filters using 100 µl of solution C6 and a centrifugation at 10,000 × g for 30 seconds.

The DNA samples was quantified using the quantifluor dsDNA system (Promega) and adjusted to obtain a concentrations with a minimum of 100 ng/µl. Sequencing was performed by the Australian Genome Research Facility (AGRF). The genomic DNA samples were amplified using several primers, including AGAGTTTGATCMTGGCTCAG and GWATTACCGCGGCKGCTG, targeting for bacterial 16S: 27F - 519R (V1-V3), CCTAYGGGRBGCASCAG and GGACTACNNGGGTATCTAAT, targeting for archaeal 16S: 341F-806R (V3-V4), CTTGGTCATTTAGAGGAAGTAA and TCCTCCGCTTATTGATATGC, targeting for fungal ITS: 1F-2R and CTGGTTGATCCTGCCAG and ACCAGACTTGCCCTCC, targeting for algal 18S: Euk3191F-EukBR. The PCR products obtained from PCR reactions were measured using fluorometry and qPCR before sequencing on the GS-FLX platform. For this procedure the resulting amplicons were denatured to give single-stranded templates. These templates were mixed with DNA Beads containing polymerase enzyme on a PicoTiterPlate device and all beads were covered with another set of *Beads* containing luciferase and sulfurylase. This device was fixed into the Genome Sequencer FLX Instrument to sequence the DNA fragments using the manufacturer's procedures.

3.1.3.3 Analysis of data

Using QIIME (Quantitative Insight Into Microbial Ecology) (Caporaso et al., 2010) the raw data were trimmed to remove uninformative sequences such as primers, barcodes, short length, chimera, homopolymers, ambiguous sequences and duplicate sequences. The reads were annotated by an online programme referred to as Metagenome Rapid Annotation using Subsystem Technology (MG-RAST) (Glass and Meyer, 2011). Using the Ribosomal Data Project II (RDP II) annotation source and recommended parameters for analysis of microbial diversity, including a minimum percentage identity cut off of 97%, an E-value cutoff of 1e-5, a minimum alignment length cutoff of 50 bp, the data obtained from 16S rRNA gene sequencing was then analysed to identify the microbial diversity of each layer and to compare the differences in each layers. Additionally, we used the rarefaction, Principle Coordinates Analysis (PCoA) (based on Euclidean distance) and heatmap graphs provided by MG-RAST. In addition to the analysis programme provided by MG-RAST, the data were saved as an excel file for further analysis.

3.1.3.4 Nucleotide sequence accession numbers

The winogradsky metagenomic analysis sets are publicly available in the MG-RAST system as F4 (bottom), F5 (middle) and F6 (surface water) samples under project identifiers 4568675.3, 4568676.3 and 4568677.3.

3.1.4 Results

One month after preparation of the columns, three different layers were visible at different depths (dark black at the bottom, green layer in the middle and a water layer at the top) (Figure 3.1.1). Following sampling from each layer, DNA was extracted and analysed through pyrosequencing of 16S ribosomal gene amplicons.

Fig 3.1.1. The Winogradsky columns.

3.1.4.1 Sequencing data

Overall, 4529, 2597 and 3816 reads were obtained from pyrosequencing of the samples from the water, middle and bottom layers (respectively) of these columns (data not shown). However, only 4436, 2446, 3628 of these sequences (respectively) could pass the quality control. A total of 411, 1118 and 1577 different reads obtained from these layers (respectively) were predicted as rRNA reads, of which only 397, 1051 and 1482 reads, respectively, belonged to known microorganisms (data not shown). The rarefaction curves showed a saturation plateau, meaning that the sequencing effort was satisfactory enough to efficiently capture the biodiversity of that environments (data not shown). Furthermore, principle coordinates analysis revealed the level of biodiversity differences in these layers (Wilke et al., 2013).

3.1.4.2 Alpha diversity

Overall, the abundance of microorganisms at the top of column was higher than in other layers, followed by the bottom and then middle of the columns (Figure 3.1.2a). The ratio of Eukarya to total reads in the corresponding layer was significantly more than those in other layers, with the water layer containing the lowest ratio of Eukarya (Figure 3.1.2b). Based on phylogenetic analysis, a total of 20 bacterial phyla and 7 eukaryotic phyla were present in the three layers; no archaea were found in these columns (Figure 3.1.3). At this level, the ratios of each phylum of microorganisms in a layer differed from that in other layers, while *Proteobacteria* (57.4%) were the dominant organisms in the water layer. These values changed to *Cyanobacteria* (22.4%) in the middle and *Bacteroidetes* (20.8%) in the bottom (Figures 3.1.3). Furthermore, statistical analysis showed that the distribution of reads related to a special microbial order in each layer and their percentage relative to the total reads of the corresponding layer was different in the three portions of the Winogradsky column (Figure 3.1.4). A list of microbial families found in these three layers of the Winogradsky columns is shown in figure 3.1.5. Moreover, a significant portion of the microbial populations in these layers (34.6%, 36.5% and 32.5%, respectively in the water, middle and bottom layers) belonged to unclassified bacteria.

3.1.4.3 Classification of microorganisms based on their physiology

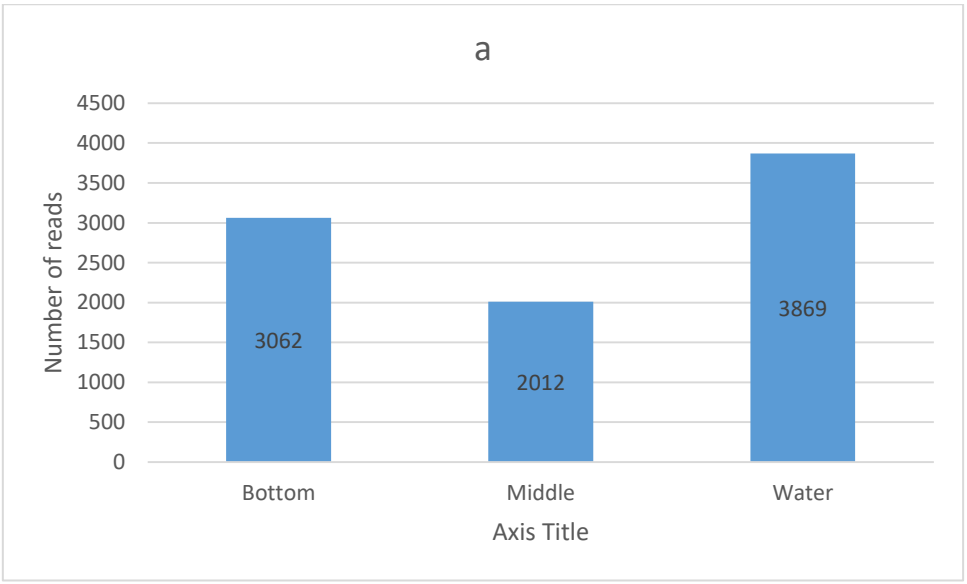
To determine the ecological relationships of the microorganisms in the three layers, we focused mainly on the ratio of known microorganisms at the order level in these three layers. Based on

the information in the literatures and microbial classification and identification handbooks (Boone et al., 2011, Whitman and Parte, 2009, Schink et al., 2011, Goodfellow et al., 2012), the known orders in this study were classified into four main metabolic groups: chemoheterotrophs, chemolithotrophs, photoautotrophs and photoheterotrophs, and the changes in these groups were analysed through comparison to the relative abundance of microorganisms in each layer and to the total number of microorganisms in the Winogradsky column. However, this classification could not be applied every time since some strains in an order may show completely different metabolic pathways from other members of that order. To cope with this, the physiology and metabolic pathways of each strain found in this study were compared with the metabolic pathways of their orders and in the case of any difference, strains were categorized separately. Based on this classification, the environmental niche of the main microorganisms in each layer of the Winogradsky column was different. Although chemoheterotrophic microorganisms were found in all three layers, the highest ratio of reads related to chemoheterotrophic microorganisms in comparison to total reads in the corresponding layer was seen at the bottom of the column (84%) (Figure 3.1.6). This ratio decreased in higher layers, falling to 18.8% at the top of this column.

The organic substrates of this column were produced by the activities of a variety of autotrophic (e.g. photoautotrophic, photoheterotrophic and chemolithotrophic) microorganisms. The majority of the autotrophic organisms (80.4%), mostly phototrophic microorganisms, inhabited the water layer of the column (Figures 3.1.6). Different phototrophic organisms were dominant in each layers of this column (*Sphingomonadales* in water layer, *Osicllatoriales* in the middle and *Bacillariophyceae* in the bottom) (Figure 3.1.7). In addition to the phototrophic organisms, a small portion (less than 1.5%) of microorganisms in each layer of the Winogradsky column was occupied by several groups of chemolithotrophic microorganisms, including sulfur oxidizing bacteria (SOBs), nitrogen oxidizing bacteria (NOBs) and hydrogen oxidizing bacteria (HOBs). The majority (80%) of the chemolithotrophic microorganisms living in the bottom of the column belonged to NOBs, followed by SOBs (13%) and HOBs (7%) (Figure 3.1.8). However, these bacteria were gradually replaced by HOBs and SOBs in the higher layers, and SOBs completely dominant SOBs in the water layer.

3.1.4.4 Classification of microorganisms based on their respiratory mechanism

A similar study was performed on the respiratory activity of microorganisms in which, based on information in the literature (Boone et al., 2011, Whitman and Parte, 2009, Schink et al., 2011, Goodfellow et al., 2012), the chemoheterotrophic bacteria found in this study were classified into aerobics, facultative anaerobes and obligatory anaerobes. Based on this classification, overall 49.0% of the microbial population at the bottom of the column consisted of aerobic microorganisms. This value increased at the top layer to 70.3% (Figure 3.1.9). The rates of obligatory anaerobic bacteria in the middle (37.5%) was higher than other layers, followed by the bottom (26.7%) and the water layer (8.3%) (Figure 3.1.9).



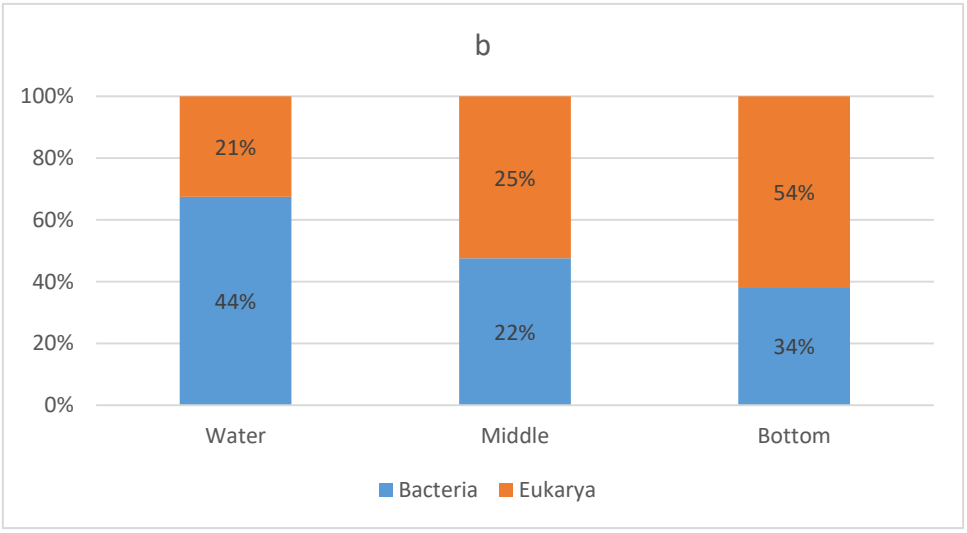


Figure 3.1.2. The abundances of read (a) and ratio of populations of Eukarya and Bacteria (b) in three layers of the Winogradsky column

Figure 3.1.3. The ratios of different phyla of microorganism in three layers of the Winogradsky column in comparison to the total reads obtained from the corresponding layer; the orders with lower than 1% are not considered.

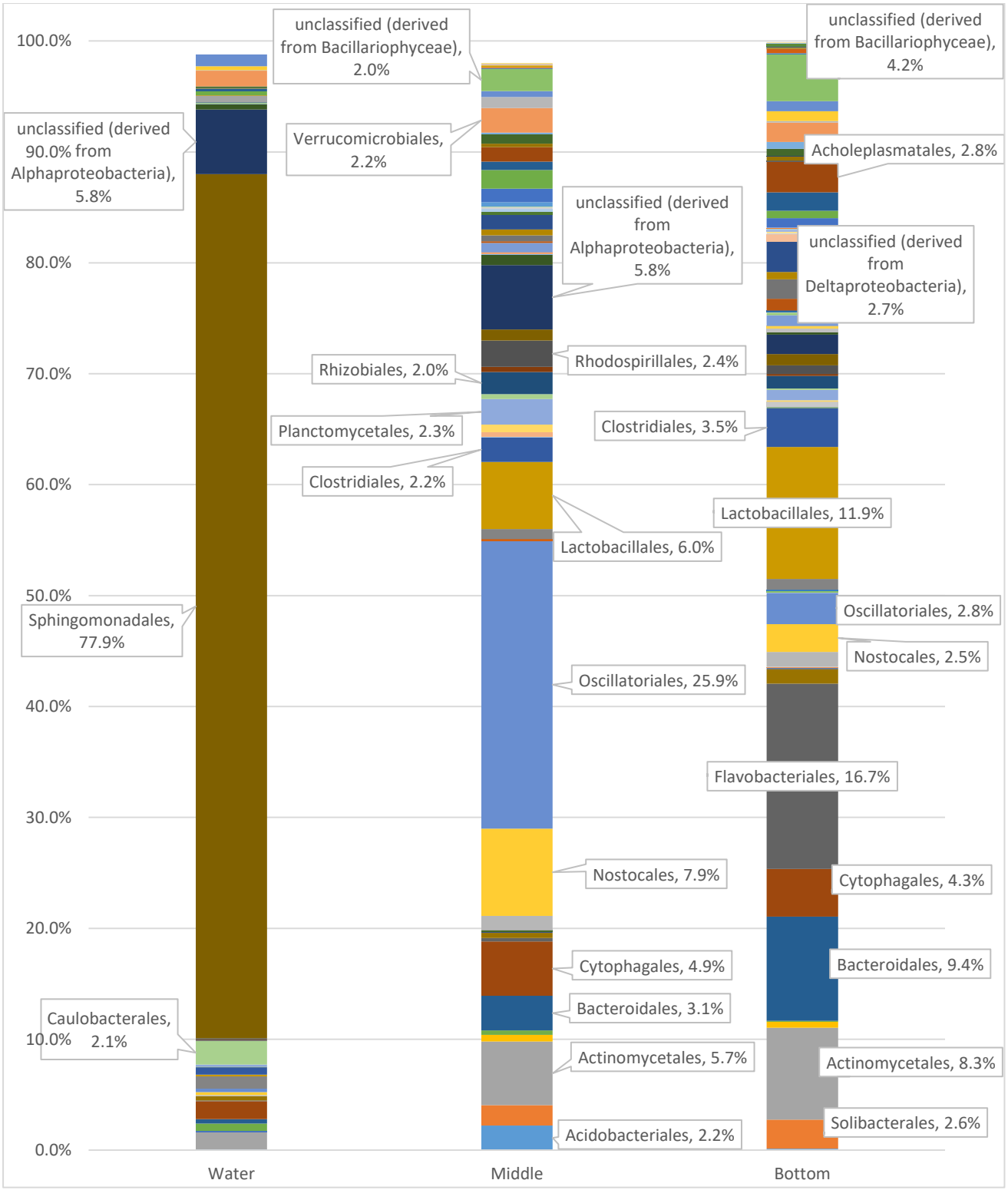


Figure 3.1.4. The ratios of order distribution in three layers (Water, middle and bottom) of column in comparison to the total microbial population in the corresponding layer; the orders with lower than 2.0% are not considered.

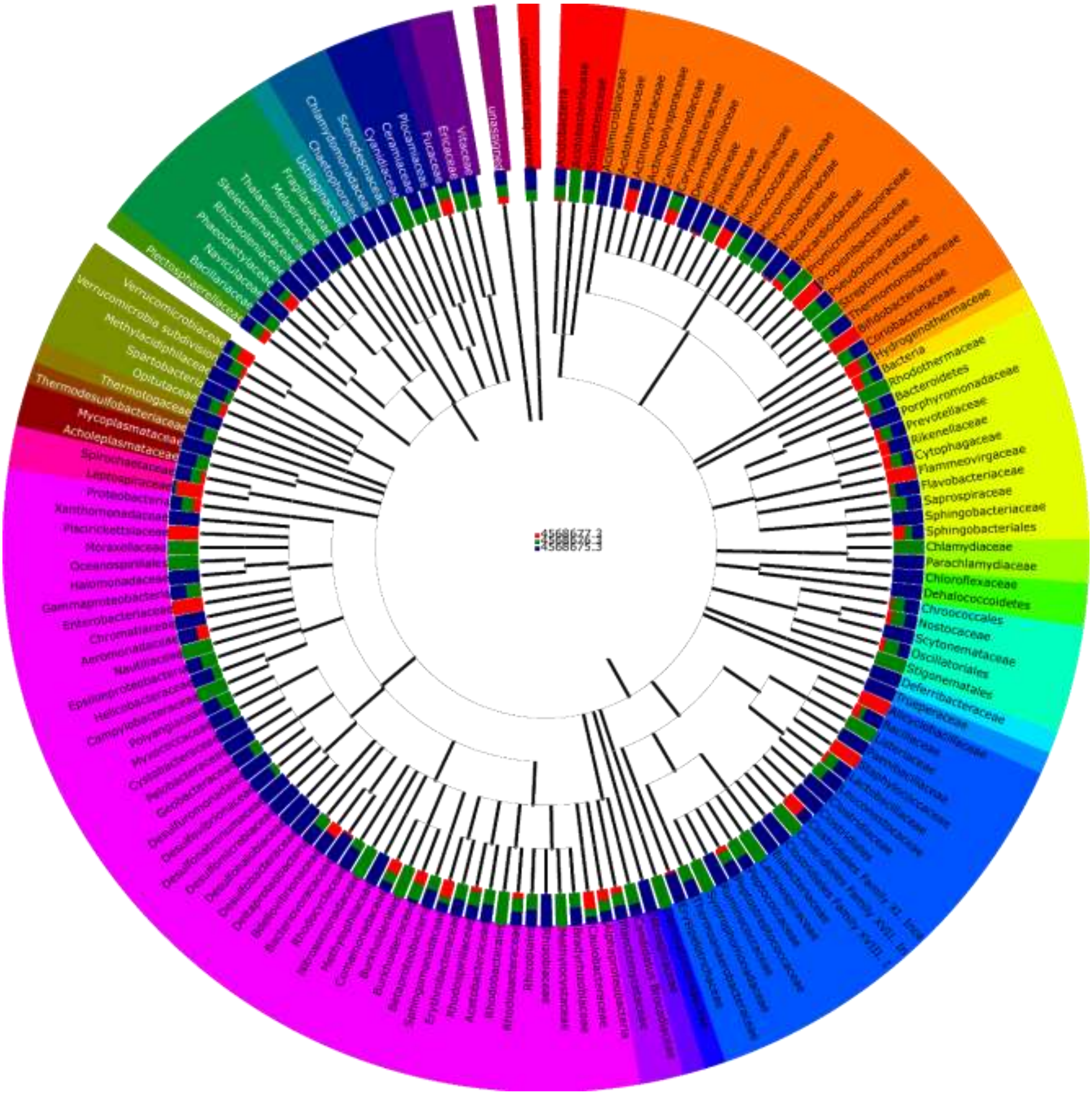


Figure 3.1.5. A list of all families found in the three layers of the Winogradsky column. Presence of each strain in a special layer is shown by three colours (Blue: bottom, Green: middle and Red: water)

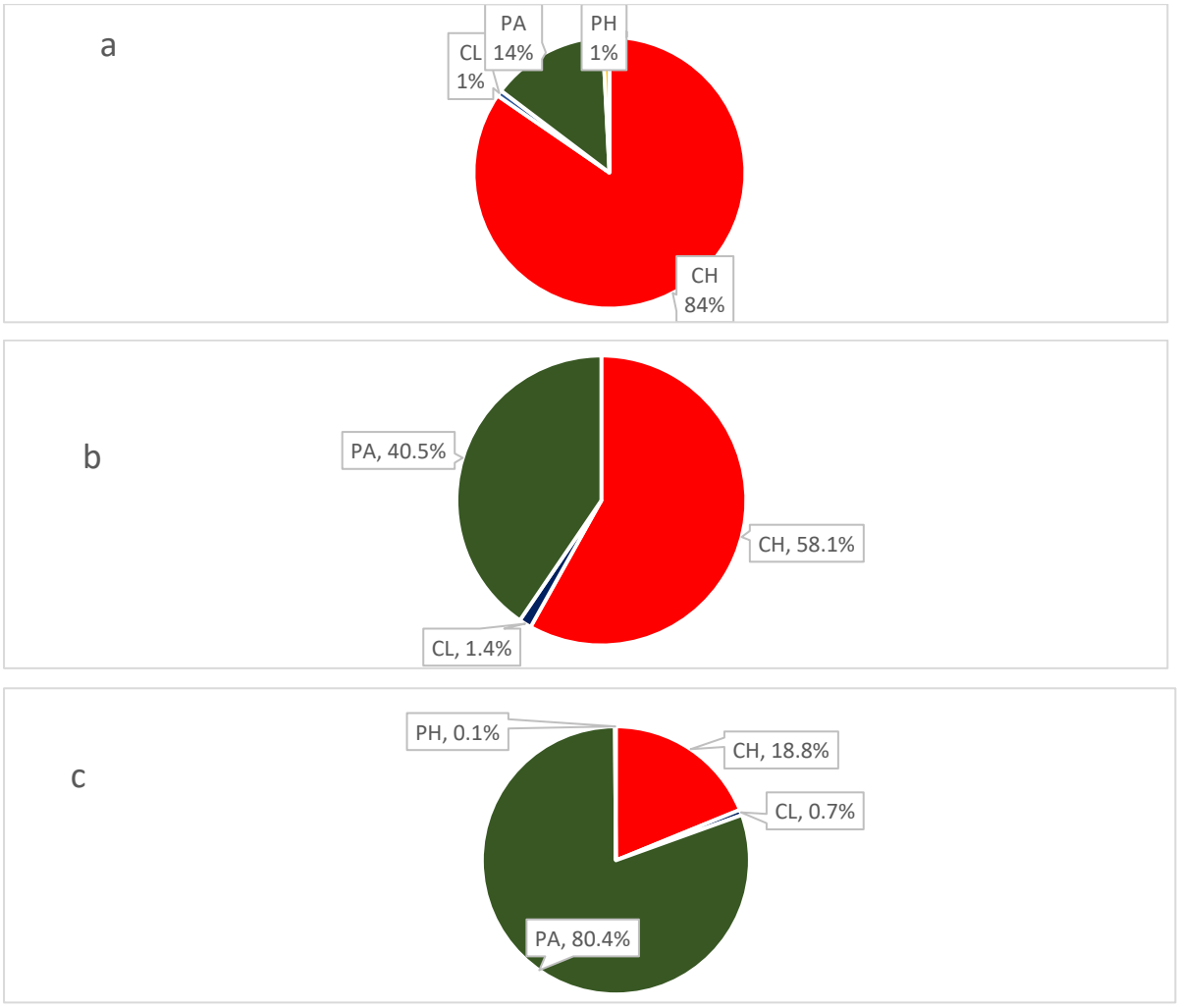


Figure 3.1.6. The ratios of microorganisms in three layers of the Winogradsky column based on their metabolic systems in comparison to total microbial population in the Winogradsky column. a: bottom; b: middle and c: water. CH: chemoheterotrophic organisms, PA: Photoautotrophic organisms, PH: photoheterotrophic organisms, CL: chemolithotrophic organisms.

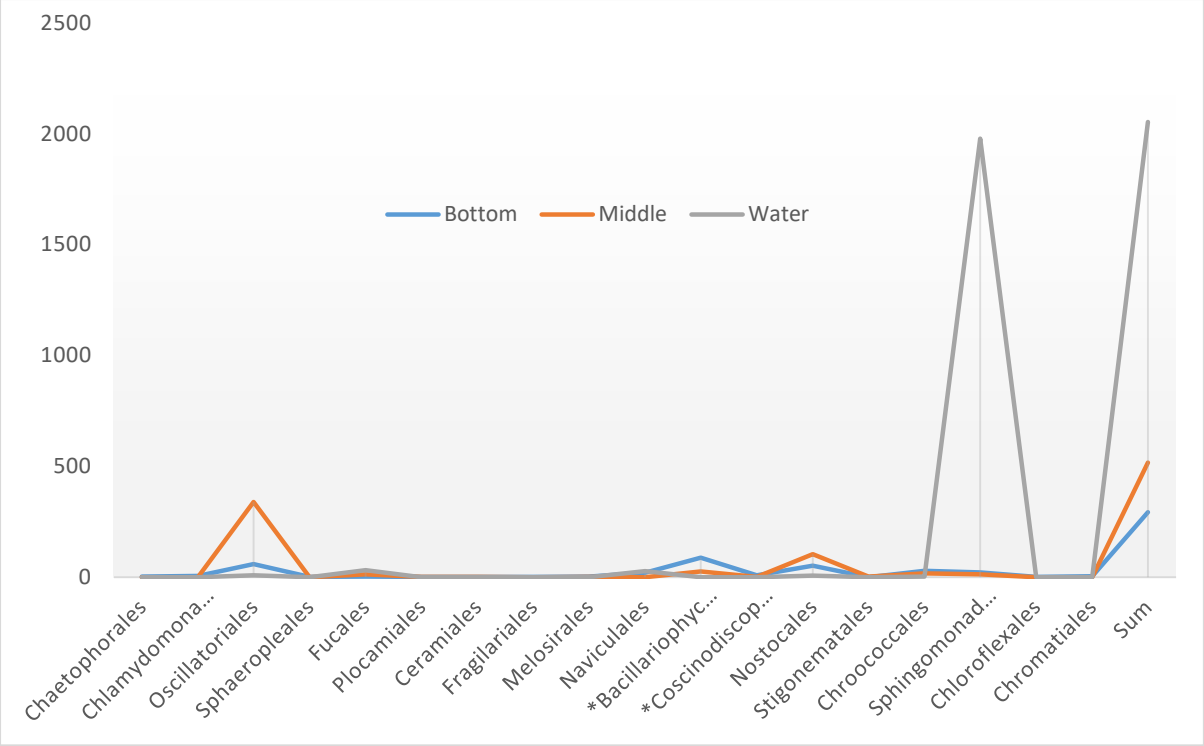


Figure 3.1.7. The abundance of photosynthetic organisms in the three layers of the Winogradsky column.

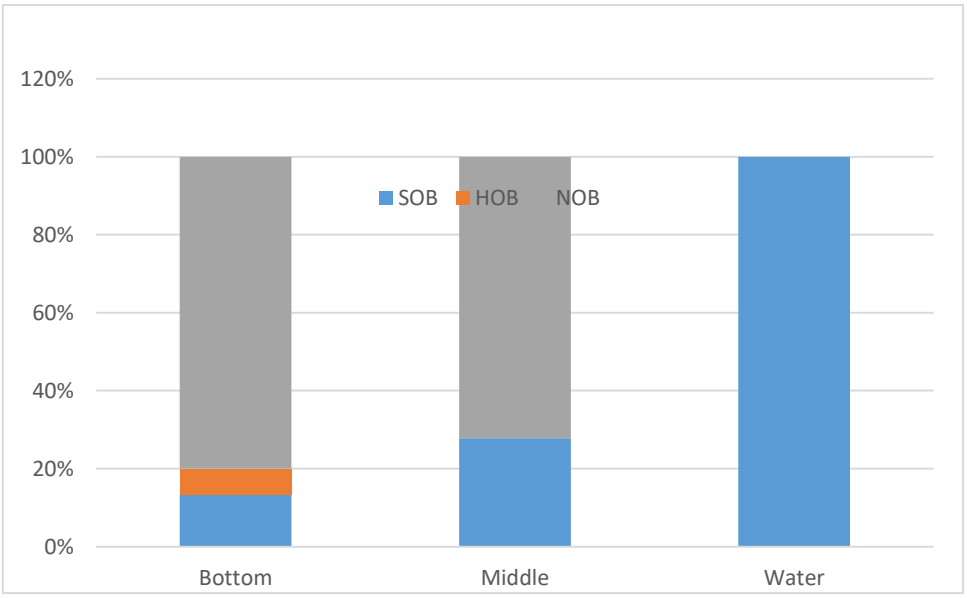


Figure 3.1.8. The ratios of chemolithotrophic organisms in three layers of the Winogradsky column in comparison to the total microbial population in the corresponding layer. SOB (Sulfur oxidizing bacteria); HOB (Hydrogen Oxidizing bacteria); NOB (Nitrogen Oxidizing bacteria).

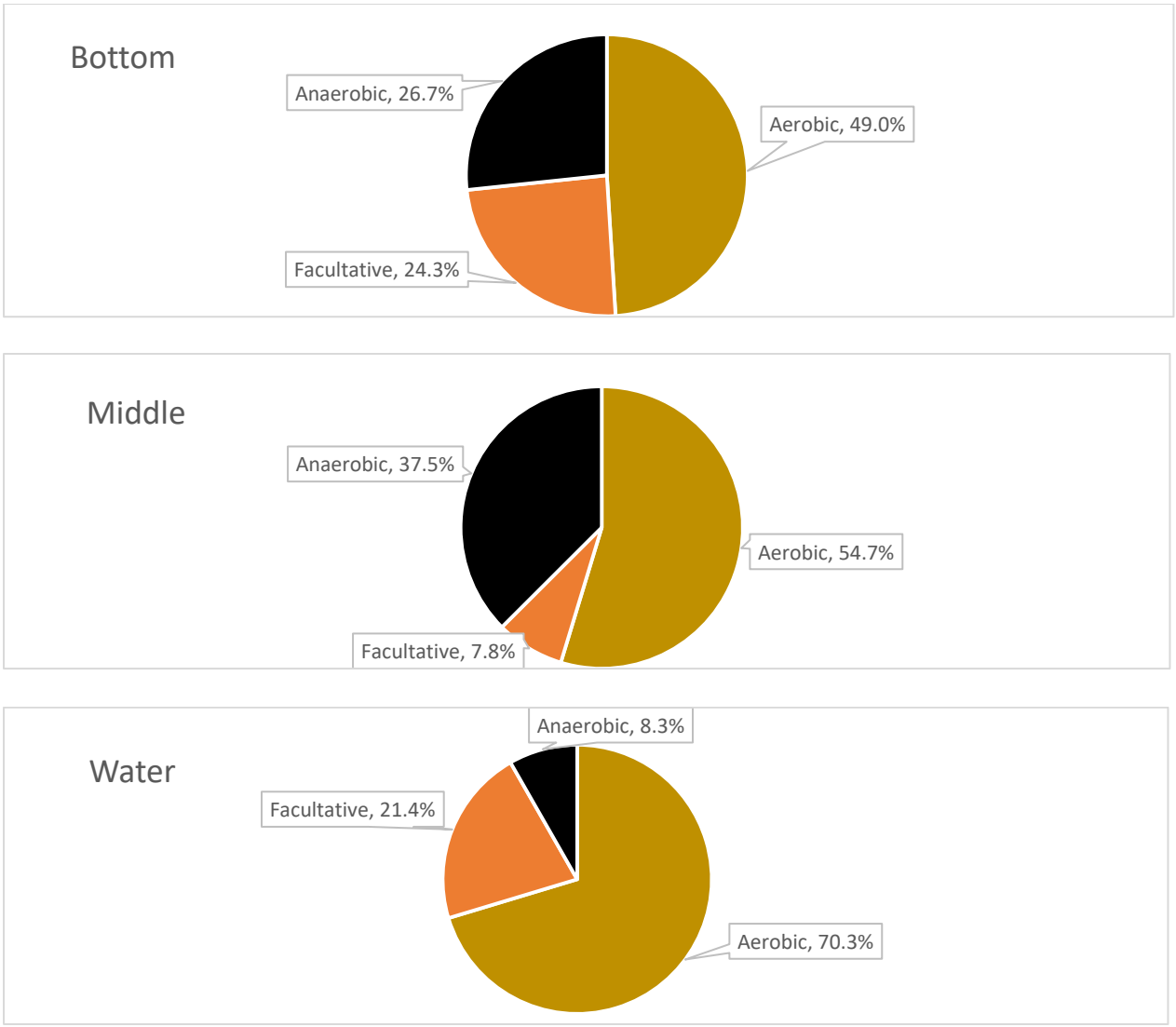


Figure 3.1.9. The ratios of heterotrophic microorganisms in three layers (bottom, middle and water) of the Winogradsky column in comparison to the total microbial population in the corresponding layer, based on their respiratory mode for oxidation of organic substrates.

3.1.5 Discussion

To determine the diversity of biological communities in a Winogradsky column and the ecological relationships governing the different layers of this column, a comprehensive study was performed on three different layers of this column using 16S rRNA gene analysis. This pyrosequencing study showed the presence of at least 27 phyla in these columns, most (20 phyla) belonged to bacteria, which overall was a little lower than the 30 phyla reported previously (Rundell et al., 2014). Although the numbers of reads obtained from the top (water) layer was more than other layers, the bottom was more diverse than other layers at the phylum

level with 24 different microbial phyla from both prokaryotic and eukaryotic domains. The complexity of biological communities decreased in the upper layers, falling to 19 and 15 phyla in the middle and water layers, respectively. These changes in the biological diversity and differences in the ratio of biological populations were observed at lower phylogenetic levels as well, ranging from 39 orders at the bottom to 33 orders in the middle and only 21 orders in the water layer. Overall, a highly complex microbial diversity in an environment indicates the presence of suitable conditions for microbial growth and activity, which in case of the Winogradsky column can be due to production of organic compounds as a result of both photoautotrophic activity and the degradation of cellulose in that environment. Although it was not very powerful factor, the impact of addition of cellulose on microbial diversity on different layers of the Winogradsky was proved in a previous study performed by Rundle et al (Rundell et al., 2014).

The surface water layer was mainly dominated by several orders of *Alphaproteobacteria*, including *Sphingomonadales* (roughly 78%) (mostly *Sandarakinorhabdus limnophila*), an unclassified bacterial order derived from *Alphaproteobacteria* (roughly 6%) and *Caulobacterales* (roughly 2%), followed by *Cytophagales* (2%), *Actinomycetales* (1.5%) and many others with less than 1.5% abundance. Therefore, *Alphaproteobacteria*, especially *Sphingomonadales*, can be used as the biomarker for the surface layer of these columns. Although the genera found in this experiment were different from those reported by Rundle et al (Rundell et al., 2014), both studies showed a dominance of *Alphaproteobacteria* at the surface water of the Winogradsky columns.

Based on the information in authorised literatures for each microbial strain, it is possible to determine the microbial niche in these three layers of the Winogradsky columns. The microorganism harbouring chlorophyll in their cells were classified as phototrophic organisms, which in turn can be divided into photoautotrophic, photolithotrophic and photoheterotrophic microorganisms based on the nature of electron and hydrogen donors used for photosynthesis (H_2O, H_2S and an organic compound, respectively) (Green and Parson, 2003). Other strains fall into the chemoheterotrophic group of organisms with the ability to use organic compounds as the sources of energy and carbon (Paul, 2014). A small narrow spectrum of microbial population referred to as chemolithotrophic microorganisms are able to use a non-phototrophic pathway to fix CO_2 at the expense of the oxidation of different mineral compounds, including

different redox forms of Fe, sulfur, hydrogen and so on (Paul, 2014). It is worthwhile to mention that although some strains of the order *Sphingomonadales* were classified as heterotrophic bacteria, the strains present in this layer, *Sandarakinorhabdus limnophila* (the most dominant species) and *Blastomonas natatoria* contain chlorophyll a and therefore are classified as phototrophic bacteria (Hiraishi et al., 2000, Gich and Overmann, 2006). In addition to these species, there were a variety of phototrophic organisms, including diatom (*Naviculales*), cyanobacteria (*Chroococcales*), green algae (*Oscillatoriales*), brown algae (*Fucales*) and purple non-sulfur bacteria (*Rhodospirillales*). However, no sulfur phototrophic bacteria have been detected in this layer probably because of shortage of H_2S in this layer. Furthermore, the water layer was occupied by only one species of chemolithotrophic bacteria *Persephonella guaymasensis*, belong to order *Aquificales*, with the ability to use both hydrogen and sulfur as the sole source of electrons to fix CO_2. As a result of this microbial combination, the majority (approximately 81%) of the abundance of microorganisms found in the water layer belonged to autotrophic organisms.

The rest of the microorganisms in this layer belonged to chemoheterotrophic bacteria, and the majority (around 70%) of these bacteria belonged to obligatory aerobic bacteria. The ratio of obligatory aerobic bacteria in this layer was significantly higher than other layers, which is expected because of dominancy of photoautotrophic organisms and the production of oxygen by these photoautotrophic organisms. No SRBs was detected in this layer due to oxygen availability and lack of sulphate.

These orders were replaced with *Oscillatoriales* (26%) (mostly *Oscillatoria spongeliae*, *Microcoleus* sp. PCC 8701 and *Oscillatoria* sp. CCAP 1459/26), *Nostocales* (8%) (mostly *Nostoc* sp. PCC 7906), *Lactobacillales* (6%), an unclassified order derived from *Alphaproteobacteria* (6%), *Actinomycetales* (6%), *Cytophagales* (5%), *Rhodospirillales* (2.4%), *Planctomycetales* (2%), *Clostridiales* (2%), *Verrucomicrobiales* (2%) and many others with less than 2% abundance in the middle of the column. These changes were associated with decreases in the ratio of photoautotrophic organisms in comparison to the water layer because of the decrease in the availability of light, and probably due to increase in the concentration of sulfur-containing compounds. The majority (40.5%) of photoautotrophic organisms in this layer belonged to different photoautotrophic microorganisms, including cyanobacteria, diatoms, brown algae, and red algae, ordered by their ratio in this layer. The dominancy of cyanobacteria in this layer can be due to employment of phycobilisomes, which contain

phycocyanin and phycoerythrin and are able to absorb the long wavelengths (red-yellow) of the light spectrum. The ratio of *Sphingomonadales* decreased in this layer probably because of inability to produce these pigments. In addition to the photoautotrophic organisms, a minor part of the autotrophic organisms belonged to chemolithotrophic organisms, including NOBs, such as orders *Candidatus Brocadiales* (*Candidatus Kuenenia stuttgartiensis*), *Nautiliales* (*Lebetimonas acidiphila*) and *Nitrosomonadales* (*Nitrosomonadaceae bacterium*) as well as a group of SOBs called as *Persephonella* sp. (belonging to order *Aquificales*).

The presence of organic compounds originated from both cellulose and the autotrophic activity of microorganisms could provide more suitable conditions for the growth of chemoheterotrophic microorganisms in comparison to the water layer. Therefore, the ratio of chemoheterotrophic microorganism in the middle of the column (58.1%) was significantly higher than this value in the water layer. Although it consists of only a small portion of the microbial community, presence of sulphate salts led to the growth of different (chemoheterotrophic) SRBs, including *Desulfovibrionales* and *Desulfuromonadales,* in this layer. Presence of this sulphate reducing orders and many other anaerobic bacteria such as *Lactobacillales*, *Planctomycetales*, *Clostridiales*, *Verrucomicrobiales* etc. increased the ratio of anaerobic bacteria in this layer. The decreases in the ratios of aerobic chemoheterotrophic bacteria in this layer occurred probably due to drop in the production of molecular oxygen, which in turn occurred because of decrease in the ratio of photoautotrophic organisms.

The light accessibility to the bottom of the column is limited only to the short wavelengths, green, blue and violet, which can be absorbed by carotenoid pigments and the phycoerythrin (found in red and blue-green algae) (Blankenship, 2013, Mishra et al., 2012). Although this spectrum of light affects differently on the growth and the levels of photosynthetic activity of organisms (Garcia-Mendoza and Ocampo-Alvarez, 2011, Mouget et al., 2004), several photosynthetic organisms, including diatoms (Gabriel Correa-Reyes et al., 2001), cyanobacteria (Garcia-Pichel and Castenholz, 1991), purple sulfur bacteria (Pfennig, 1967) are able to use this spectrum for CO_2 Fixation. The phototrophic producers in bottoms of these columns (15% of the total population of this layer) belonged to different photoautotrophic categories, including diatoms (including *Bacillariophyceae*, *Naviculaceae*, *Phaeodactylaceae*, *Coscinodiscophyceae*, *Melosiraceae* and *Fragilariaceae*), cyanobacteria (*Nostocaceae*, *Scytonemataceae* and unclassified *Chroococcales*, unclassified *Oscillatoriales*), green algae (*Chlamydomonadaceae*, unclassified *Chaetophorales*, *Scenedesmaceae*), purple non-sulfur

bacteria (*Rhodospirillaceae* and *Acetobacteraceae*), purple sulfur bacteria (*Chromatiaceae*), red algae (*Ceramiaceae* and *Plocamiaceae*), brown algae (*Fucaceae*) and green non-sulfur bacteria (*Chloroflexaceae*) (ordered here by ratio from highest to lowest), which are able to absorb the short wavelength light.

Furthermore, the heterotrophic life at the bottom is supported by other autotrophic sources produced by chemolithotrophic bacteria (1%), mostly ammonia oxidizing bacteria such as *Nitrospirales* (*Leptospirillum ferrooxidans*) and *Candidatus Brocadiales* (*Candidatus Kuenenia stuttgartiensis*) and a group of SOBs derived from *Aquificales* (*Persephonella* sp.) and a group of HOBs derived from *Dehalococcoidetes* (*Dehalococcoides* sp. TM-EtOH) which, based on the microorganism, use different organic and/or inorganic materials such as sulfur, ammonium, nitrite and hydrogen to fix CO_2. These autotrophic organisms, as well as the cellulose present at the bottom and the downward movement of organic compounds produced in the upper layers provide the organic compounds for chemoheterotrophic organisms in this layer.

The chemoheterotrophic bacteria comprised 84% of the total microbial population in these layers, mostly belonging to *Flavobacteriales* (roughly 16.7%) (mostly *Flavobacterium columnare*), *Lactobacillales* (roughly 12%) (mostly *Lactobacillus casei* ATCC 334, *Lactobacillus paracasei*, *Leuconostoc mesenteroides* and *Pediococcus pentosaceus*), *Actimomycetales* (8.3%) (mostly *Nocardioides* sp. MTD22), *Bacterioidales* (mostly *Porphyromonas endodontalis*, *Tannerella forsythia* and *Prevotella buccae*) (6.3%), *Cytophagales* (4.3%) (such as *Cytophaga* sp., *Flexibacter litoralis*, *Flexibacteraceae bacterium* VUG-A141a, *Marinoscillum furvescens*), an unclassified organism derived from *Bacillariophyceae* (4.2%), Clostridiales (3.5%) (such as *Clostridium* sp., *Tepidimicrobium ferriphilum*, *Helcococcus kunzii*, *Sedimentibacter hydroxybenzoicus*, *Symbiobacterium thermophilum*, *Eubacterium* sp., *Desulfitobacterium hafniense*, *Pelotomaculum propionicicum*, *Acetivibrio cellulolyticus*), an unclassified order derived from *Bacteroidetes* (3.5%) (*Candidatus* sp., and *Prolixibacter bellariivorans*) many orders being less than 3% population. Among these chemoheterotrophic bacteria, roughly 51% belonged to obligatory and facultative anaerobic bacteria. Thereby, the dominant bacteria at the bottom are characteristic of a more anaerobic environment in which chemoheterotrophic bacteria use anaerobic respiration or fermentative pathways to catabolize organic compounds to provide their energy and carbon sources. Among the anaerobic bacteria, the sulphate reducing bacteria

in this layer, including *Desulfobacterales*, *Desulfovibrionales* and *Desulfuromonadales*, contain 3% of the total abundance of the microbial population in this layer. Although no oxygen is penetrating to the bottom of this column from outside, the presence and activity of photoautotrophic microorganisms produce the oxygen required for growth of aerobic microorganisms, which consist 49% of the heterotrophic population in the bottom.

It is worthwhile to remark that the bottoms were the only part of the Winogradsky columns where a group of anoxygenic photolithotrophic organisms, *Chromatiaceae*, and two groups of photoheterotrophic bacteria, *Rhodosiprillales* and *Chloroflexaceae*, were found, certainly because of presence of H_2S and their ability to capture the light. Furthermore, some groups of microbial population, including *Truepera radiovictrix* DSM 17093 (an aerobic chemoheterotrophic bacterium belonging to phylum *Deinococcus-Thermus*) (Albuquerque et al., 2005), *Kosmotoga olearia* TBF 19.5.1 (an anaerobic chemoheterotrophic bacterium belonging to *Thermotogae*) (Whitman and Parte, 2009), *Deferribacter desulfuricans* SSM1 (a sulfur-, nitrate and arsenate reducing bacterium belonging to phylum *Deferribacteres* (Takai et al., 2003) and *Leptospirillum ferrooxidans* (an acidophilic iron oxidizing bacterium belonging to *Nitrospirae*) (Slobodkina et al., 2009, Garrity and Holt, 2001) and *Roseiflexus castenholzii* and *Dehalococcoides* sp. TM-EtOH (two autotrophic non-sulfur bacteria belonging to phylum *Chloroflexi*) (Whitman and Parte, 2009) were unique to the bottom of the winogradsky column.

Overall, this large variation in the diversity of populations provides evidence for the presence of three different niches in these layers, caused by the concentration of sulphate and availability of light in these layers. The water layer was the most abundant part of the Winogradsky column, dominated mostly by photoautotrophic organisms and aerobic chemoheterotrophic bacteria. However, the bottom contains the most microbial diversity in this column, mostly dominated by anaerobic chemoheterotrophic organisms. While the ratio of Alphaproteobacteria in the water layer was significantly higher than for the bottom layer, the ratios of several phyla, including *Bacteroidetes*, *Firmicutes*, *Nitrospirae*, *Spirochaetes*, *Tenericutes*, *Thermodesulfobacteria*, *Verrucomicrobia* decreased in middle and water layers in comparison to the bottom. Therefore, the presence of these microorganisms can be used as a biomarker in the bottom. These changes in microbial diversity and their ratios in different layers of the Winogradsky columns indicates the presence of different environments in these column in which the ratio of phototrophic and aerobic chemoheterotrophic organisms decreased from the water to bottom. Although the diversity and ratio of microorganisms determined in this study

was quite different from that in previous study (Rundell et al., 2014), the rates of decreases in the gradient of photoautotrophic organisms and aerobic chemoheterotrophic bacteria from water layer to bottom were similar. The variability in the microbial diversity and their ratios in different studies can be because of the original sediment samples used for preparation of the columns, the heights of the columns, the incubation temperature, light wavelength and incubation period. Therefore, these factors must be considered in case of design of future studies to determine effects of a factor, such as chemical contaminants and physical factors, on total or partial microbial population in a sediment.

3.2 Effects of crude oil contamination on microbial diversity in the Winogradsky column as a fresh water lake biome

Submitted to: Journal of Environmental health sciences and Engineering

3.2.1 Abstract

Background: Crude oil contamination in the natural environments has been considered as a serious concern due to its adverse effects on biota including different microorganisms. However, no comprehensive study has been performed using new emerged technologies to study the effects of crude oil contamination on total microbial population in different environments. In this study, a pyrosequencing approach was employed to investigate the impacts of different concentrations (3%, 1%, 0.5% and 0.1%) of crude oil contamination in the winogradsky column (microcosm), which mimics a fresh water lake biome.

Results: Based on the results, crude oil pollution enhanced the microbial diversity and their ratios in different layers of this microcosm, mostly belonged to chemoheterotrophic prokaryotes. Furthermore, there were increases in the ratio of aerobic bacteria in the water layer and middle of the microcosm. However, the ratio of oxygenic photoautotrophic organisms decreased in all layers of the contaminated column and the effect was more pronounced with an increase in oil concentration.

Conclusion: Crude oil contamination led to intensive changes in the diversity of microorganisms of sediments, which caused alterations in the environmental niche of the given area.

3.2.2 Introduction

Environmental contamination through the release of crude oil and its products is a global environmental issue, and this concern is increasing as a result of population growth and industrial expansions (Hao et al., 2004). Oil spillage into the environment can be happened through different activities in the oil industries, including exploration, extraction, transportation, refining and oily waste management. Furthermore, huge oil spills are occasionally occur as a result of subversions on pipelines and tanker accidents. The release of

2-3,000,000 US barrels crude oil into the Persian Gulf War in 1991(Ross, 1991) and 350 tonnes heavy oil into the New Zealand coast through a ship accident in 2011(Hunt et al., 2014), are two examples of environmental disasters. Crude oil consists of a mixture of various hydrocarbons, including aliphatics and aromatics, which are categorized as serious and hazardous environmental pollutants (Nolvak et al., 2012, Anthony, 2006). However, the degree of toxicity of crude oil and its products depends on its chemical composition and their exposure concentration in the environment (Hao et al., 2004, Sathishkumar et al., 2008). While several microorganisms are directly affected by crude oil and/or its products, neutral or positive relationships between microorganisms and oil constituents have been proved in many environments (Anthony, 2006). In the positive interactions, microorganisms are able to use these molecules as their carbon and energy sources (Nolvak et al., 2012, Boopathy, 2000). Although oil-degrading microbes are ubiquitous and able to consume oil compounds, their oil contaminants tolerance level and oil degrading capacity are varied (Nolvak et al., 2012, Boopathy, 2000).

On the other hand, there are high levels of positive and negative relationships between metabolic activities of heterotrophic and autotrophic microorganisms and higher levels of life (Kirchman, 2010). Therefore, any changes in the microbial community may impact the whole nutritional networks in this ecological system (Treseder et al., 2012, Orcutt et al., 2011). In addition, since microbial activity is critically necessary for oil bioremediation, investigation on the microbial diversity in polluted environments is beneficial to identify the active members of an oil utilising community and the changes in response to environmental parameters and specific oil substrates (Mail et al., 2012). This information finally assists scientists to set up chemical, physical and environmental conditions that may improve biodegradation of oil and its products.

Although microbial diversity is traditionally characterized using culture based techniques, these methods are limited by the fact that many oil degrading organisms do not grow on defined microbial media (Head et al., 2006, Mail et al., 2012). Furthermore, application of traditional random “shotgun” sequencing using Sanger DNA sequencing (dye terminator sequencing) is limited to the number of clones chosen accidentally for sequencing (Abbasian et al., 2015b). Pyrosequencing approaches, however, are theoriticallay able to track the microbial diversity

and their ratios in a complex community without these limitations (Abbasian et al., 2015b). Therefore, in this study we used pyrosequencing approach to characterize the effects of different concentrations of crude oil on the microbial diversity in a Winogradsky column-based microcosm.

3.2.3 Material and methods

3.2.3.1 Preparation of the Winogradsky column and oil spiking

The sediment samples were collected from a local lake located at Mawson lakes, Adelaide South Australia. The Winogradsky columns were prepared in ten plastic transparent columns (9 cm in diameter and 25 cm high). Following sieving the sediments to remove any macroscopic animals, roots and rubbles, one part of the sediment was spared for preparation of the bottom of the columns, in which 0.50% W/W Na_2SO_4, 0.25% W/W $CaCO_3$ and around 20 g of shredded newspaper, as source of cellulose, were added to provoke cellulolytic anaerobic bacteria. This combination was used to fill the first third (bottom) of each column. The second third (middle) of these columns was filled with the sieved sediments without any nutritional supports. The upper layer of the column was divided into two, in which only one part (one sixth) was filled with the lake water. The columns were sealed with nylon to prevent any air exchanges with outside. Furthermore, two thirds of the column from the bottom was roofed with aluminium foil to avoid light irradiation from the sides. The columns were kept in a phycology laboratory under a constant exposure to light (400–700 nm wavelength), at a temperature of 25° C for 30 days before spiking with oil. After the incubation period, the columns were amended with different concentrations of crude oil (3%, 1%, 0.5% and 0.1%). Oil was added from the top of the columns in order to mimic a crude oil spillage in a fresh water lake. No oil was added to the control columns. Two sets of each control column and the columns treated with different concentrations of crude oil were prepared in this study.

3.2.3.2. TPH analysis in the crude oil contaminated sediments

The sequential ultrasonic solvent extraction method (Risdon et al., 2008, Ramadass et al., 2015) was used to extract TPH compounds from the soil sediment samples. Quantification of TPH concentration was carried out using a gas chromatograph fitted with a flame ionization detector (GC-FID, Agilent 6890). Chromatography was performed on a fused-silica capillary column

BPX-5 from SGE (15 m × 0.32 mm internal diameter) coated with HP-5 (0.10-μm film thickness). Helium was used as the carrier gas at 2.5 mL min^{-1}, and the FID detector temperature maintained at 300°C. Splitless injection with a sample volume of 1 μL was applied. The oven temperature was increased from 50°C to 300°C at a gradient of 25°C min^{-1} and held at this temperature for 5 min. The total run time was 19.6 min. Hydrocarbons were quantified using Agilent Chemstation Software by integration and calibration of peaks of a known concentration of an external calibration standard– Hydrocarbon Window Defining Standard (C_8-C_{40}) from AccuStandard®(Miller and Miller, 1987). Five concentrations of external calibration standard in the range expected in the samples were analysed; a linear curve fit with a R^2 value of 0.997 was obtained. The CCV (Continuing Calibration Verification) was analyzed at the start and end of every 20 samples and CCV recovery was 95-110% of true value. Hexane was run as blank with every 10 samples to demonstrate that the system was free from contamination. The surrogate (*o*-terphenyl) was spiked at a level to produce a recommended extract concentration of 20 μg/mL. Surrogate recoveries lay in the range 70-120% for all the samples analyzed. 35 mg kg^{-1} was the minimum concentration of TPH detected (MDL) through this analytical method.

3.2.3.3 DNA extraction and pyrosequencing process

Samples were collected after 30 days from three layers, the water layer at the top of the column, the 1/3 middle portion and the 1/3 bottom layer. DNA extractions were conducted by the power soil DNA isolation kit (MO BIO) and according to manufacturer's instruction. Using a quantifluor dsDNA system (Promega), the extracted DNAs were quantified to confirm100 ng DNA per microliter solution. Following an DNA amplification process using the primers targeted for 16SrRNA of bacteria (referred to as 16S: 27F/519R; 5′-AGAGTTTGATCMTGGCTCAG-3′ and 5′-GWATTACCGCGGCKGCTG-3′) (Lane, 1991) and Archaea (referred to as 16S: 341F-806R; 5′-CCTAYGGGRBGCASCAG-3′ and 5′-GGACTACNNGGGTATCTAAT3-′) (Abbasian et al., 2015c) as well as 18SrRNA of Fungi (referred to as ITS1F-2R; 5′-CTTGGTCATTTAGAGGAAGTAA-3′ and 5′-TCCTCCGCTTATTGATATGC-3′) (Hamelin et al., 1996) and algae (referred to as Euk3191F-EukBR; 5′-CTGGTTGATCCTGCCAG-3′ and 5′-ACCAGACTTGCCCTCC-3′) (Wilms et al., 2006), the PCR products were sequenced on the GS-FLX platform located at the Australian Genome Research Facility (AGRF).

3.2.3.4 Pyrosequencing data analysis

The raw data obtained from pyrosequencing platform were trimmed using QIIME (Quantitative Insight Into Microbial Ecology) (Caporaso et al., 2010) to remove any uninformative data, including duplicate sequences, ambiguous sequences, homopolymers, chimera, short length, barcodes and primers. Next, following submission of the data in Metagenome Rapid Annotation using Subsystem Technology (MG-RAST) (Glass and Meyer, 2011), the data were analysed using the Ribosomal Data Project II (RDP II) annotation source and with respect to the parameters recommended for analysis of microbial diversity (a minimum alignment length cutoff of 50 bp, an E-value cutoff of 1e-5 and a minimum percentage identity cut off of 97%) to identify the microbial diversity in each sample. To calculate the Sharon's index and the effective numbers of strains, the datasheets were transferred to Excel software.

3.2.3.5 Nucleotide sequence accession numbers

These pyrosequencing data are publicly available in the MG-RAST system as F7, F8, F9, F11, F12, F13, F14, F15, F16, F19, F20 and F21 samples under project identifiers 4568678.3, 4568679.3, 4568680.3, 4568664.3, 4568665.3, 4568666.3, 4568667.3, 4568668.3 4568669.3, 4568672.3, 4568673.3 and 4568674.3, respectively.

3.2.4 Results and discussion

Oil contamination pose serious challenges for microorganisms in both terrestrial and aquatic environments. In addition to the alternations in the environmental physico-chemical factors, microorganisms are challenged by the complexity and toxicity of oil hydrocarbons, their insolubility in water and the low bioavailability of carbon and other essential nutrients (Morgan and Watkinson, 1994, Sánchez-Palencia González, 2011). In order to determine the adverse effects of crude oil contamination on microbial diversity of a fresh water lake habitat, the Winogradsky column-based microcosms were spiked with different concentrations of crude oil, and the effects on microbial diversity were investigated by a FLX-pyrosequencing platform and bioinformatic analysis. As like as previous study performed by Abbasian and his colleagues (Abbasian et al., 2015d), each layers of the control column contained a unique

profile of microbial diversity as a result of sensitivity to H_2S and their accessibility to O_2 and light.

One month incubation period before spiking led to formation of three different layers in the columns: a black bottom, a green middle and a water layer. TPH analysis on three layers of two tested concentrations (0.1% and 3%) showed that all three layers were contaminated with crude oil throughout the incubation time, and the levels of these compounds at the bottom were higher than middle and surface water layers (Table 3.2.1). The presence of hydrocarbons in different depths of the columns could affect the microbial abundance and diversity in that habitats.

Concentration	Layer	>C10-C16 (PPM)	>C16-C34 (PPM)	>C34-C40 (PPM)	TPH (PPM)
0.1%	Surface water	43	65	6	114
	Middle	47	125	6	178
	Bottom	65	155	7	227
3%	Surface water	1248	2054	13	3315
	Middle	1320	2094	14	3428
	Bottom	1433	2592	13	4037

Table 3.2.1. The crude oil concentrations in three layers of the winogradsky column after spiking with 0.1% and 3% crude oil; PPM= Particle Per Million.

The volumes of data obtained from pyrosequencing study of the samples from these three layers of the control Winogradsky columns, and columns contaminated with 0.1%, 0.5%, 1% and 3% crude oil are mentioned in table 3.2.2. Although the colour of these three layers did not change after spiking and one month incubation thereafter, 16S rRNA analysis showed significant alterations in the microbial diversity in the layers of the Winogradsky columns after treatment. These microorganisms belonged to different phyla of bacteria and Eukarya, and no archaea was detected in these samples. Based on saturation plateaus obtained from the rarefaction curves, the numbers of individual species in these samples were representative of those present in the columns (data are not shown). Furthermore, PCoA test showed that the

samples prepared from each layer (water, middle and bottom) contained different levels of microbial diversity (data are not shown).

To start with, the total numbers of reads obtained from the water layer of control column was less than other layers in columns (Table 3.2.2). Although the total abundances of microorganism in the water layers and bottoms of these columns grew after treatment, crude oil had a negative effect on this value in the middle of the column. The treatments with crude oil impacted differently on each group of microorganisms (Fig. 3.2.1, 2 and 3), and therefore, could change the microbial diversity in this layer as was confirmed by Sharon's index (Table 3.2.3). Some orders, such as Acidobacterials, Deinococcales, Aeromonadales and Pasteurellales were completely disappeared in the water layer as a result of the lowest degrees (0.1%) of oil spiking. Furthermore, the greater levels of oil spiking, the greater decreases found in the ratios of Sphingobacteriales, Chroococcales, Rhizobiales, Nitrospirales, Nostocales and Naviculales. Overall, this study showed that several microorganisms such as Chroococcales, phototrophic Sphingomonadales and Oscillatoriales, all belonged to phototrophic organisms, are highly susceptibility to compounds of crude oil.

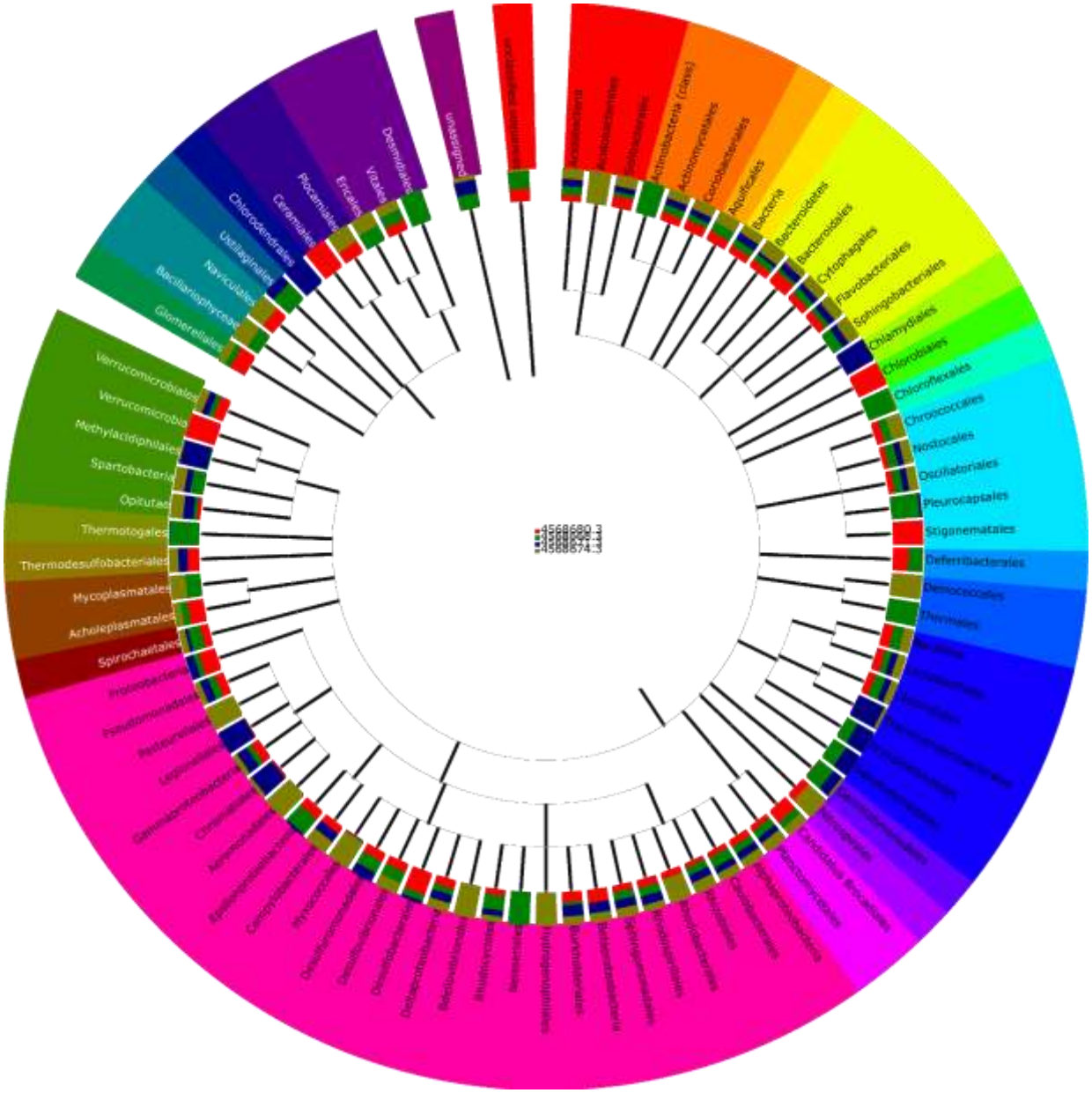


Figure 3.2.1 The microbial diversity in water layer of winogradky column contaminated with different concentrations (control, 0.5%, 1% and 3%) of crude oil. All microbial orders belonging to a phylum are indicated using different colours.

Dataset	Layer	Name	Raw number of nucleotide Count (bp)	Post QC: number of nucleotide Count (bp)	Upload: Sequences Count	Post QC: Sequences Count	Post QC: Mean Sequence Length (bp)	Processed: Predicted rRNA Features	Alignment: Identified rRNA Features
Control	Surface water	F21	1,002,129	980,544	2,042	1,972	497 ± 19	751	693
	Middle	F20	1,283,870	1,253,327	2,605	2,505	500 ± 18	897	853
	Bottom	F19	1,068,134	1,043,287	2,129	2,044	510 ± 20	1,018	945
0.1%	Surface water	F18	1,257,619	1,234,690	2,529	2,449	504 ± 21	822	801
	Middle	F17	1,068,134	1,043,287	2,129	2,044	510 ± 20	1,018	945
	Bottom	F16	1,466,147	1,442,951	2,873	2,814	512 ± 17	1,061	991
0.5%	Surface water	F15	1,346,736	1,304,918	2,755	2,614	499 ± 20	902	842
	Middle	F14	1,032,055	1,012,226	2,076	2,022	497 ± 39	733	699
	Bottom	-	-	-	-	-	-	-	-
1%	Surface water	F12	1,346,283	1,309,705	2,746	2,610	501 ± 23	973	916
	Middle	F11	993,164	956,339	2,041	1,907	501 ± 25	752	706
	Bottom	F13	824,076	806,630	1,627	1,571	513 ± 18	504	473
3%	Surface water	F9	1,382,020	1,347,608	2,798	2,670	493 ± 64	1,030	973
	Middle	F8	1,108,528	1,082,603	2,219	2,129	499 ± 54	917	873
	Bottom	F7	1,706,733	1,671,904	3,434	3,276	510 ± 25	1,225	1,149

Table 3.3.2. The volumes of data obtained from pyrosequencing study of the samples from these three layers of the control Winogradsky columns, and columns contaminated with 0.1%, 0.5%, 1% and 3% crude oil; QC: quality control.

Subject	Index	Control	0.1%	0.5%	1.0%	3.0%
Water	H	4.0	2.8	3.5	3.3	3.1
	D	52.9	16.6	32.4	28.3	23.3
Middle	H	3.2	2.8	3.5	3.5	3.4
	D	24.0	16.0	33.6	33.6	30.2
Bottom	H	3.7	3.5	-	3.8	3.7
	D	41.5	34.4	-	45.8	38.5

Table 3.2.3. The Sharon's index (H) and the effective numbers of strains (D) in three layers of the Winogradsky column after spiking with 0.1% and 3% crude oil.

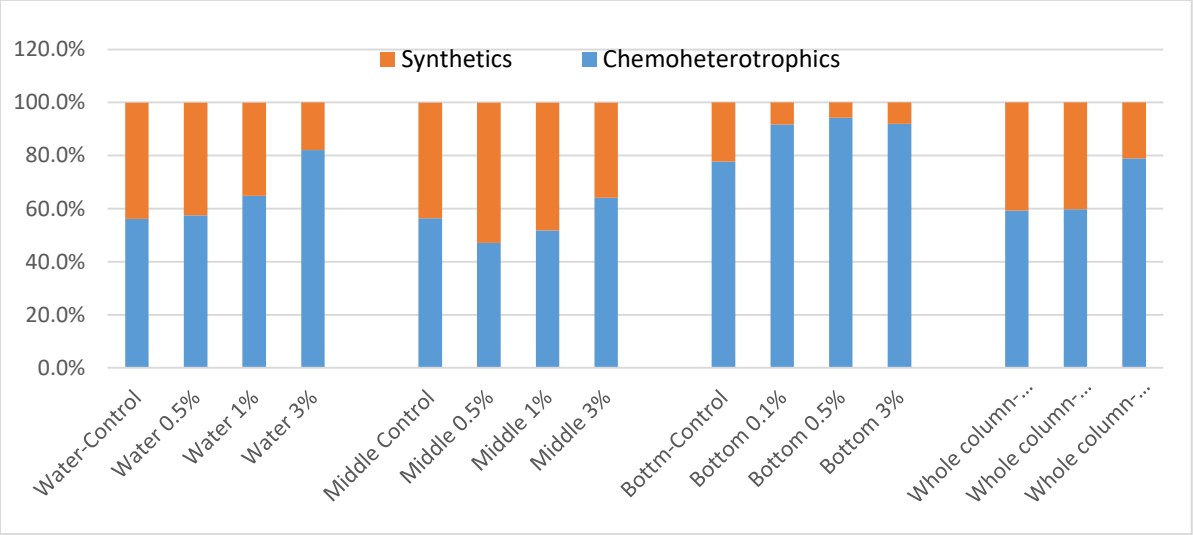


Figure 3.2.2 Ratios of chemoheterotrophic and phototrophic microorganisms present in different layers of the Winogradsky column contaminated with different concentrations of crude oil.

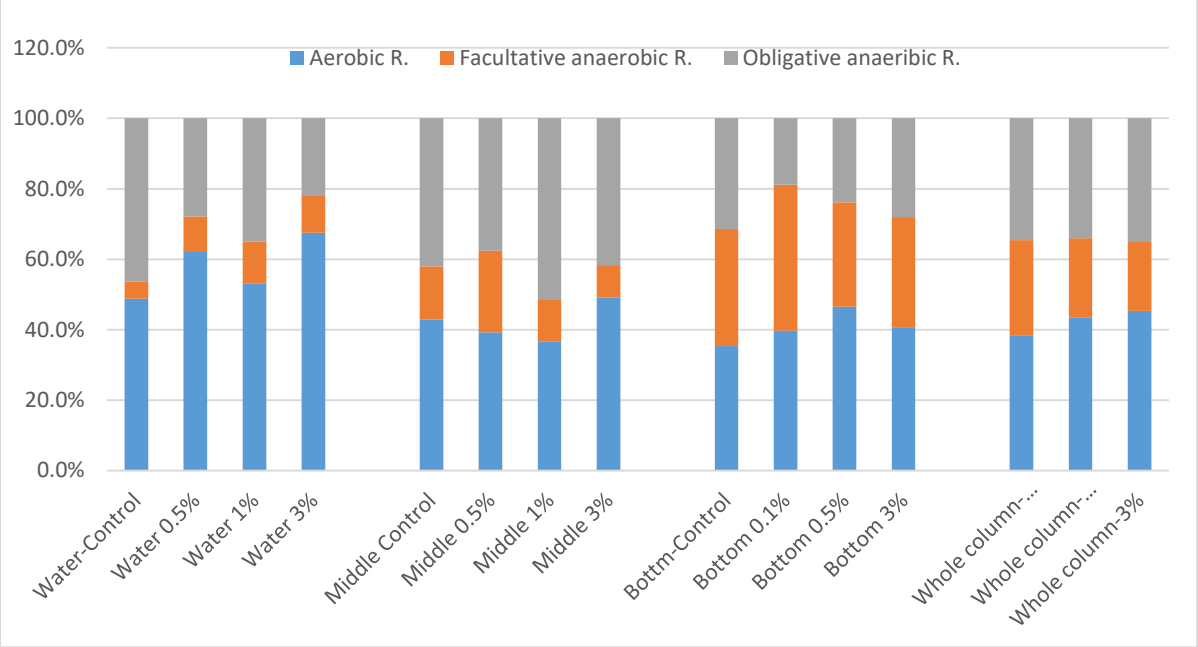


Figure 3.2.3 Ratios of chemoheterotrophic microorganisms, categorized based on their respiration activity, in different layers of the Winogradsky columns contaminated with different concentrations.

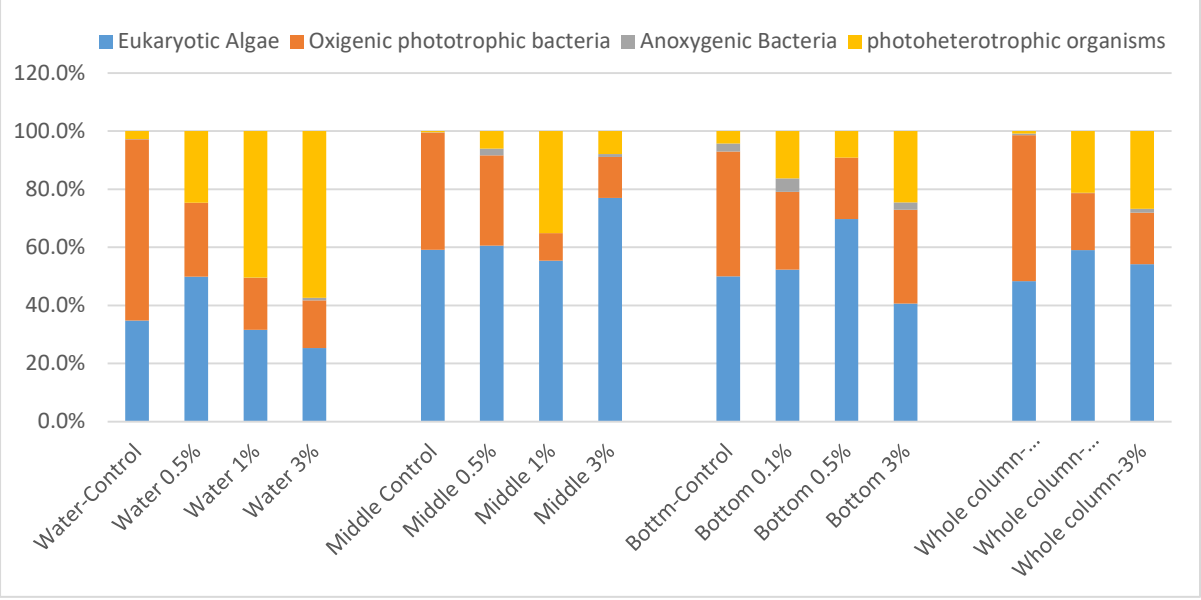


Figure 3.2.4 Ratio of phototrophic organisms in the layers of the Winogradsky columns contaminated with different concentrations of crude oil.

However, this treatment has positive effects on the ratios of other bacterial orders such as Actinomycetales (*Corynebacterium* sp., *Frankia* sp., *Geodermatophilus* sp., *Microbacterium* sp., *Arthrobacter* sp., *Micromonospora* sp., *Mycobacterium* sp., *Nocardioises* sp. and *thermomonospora* sp.), Clostridiales (mainly *Fusibacter* sp., and *Clostridium* sp.), Bacillales (mainly *Baillus* sp.) and Lactobacillales (*Lactobacillus* sp., *Leuconostoc* sp. and *Pediococcus* sp.), all belonged to gram positive bacteria, in all layers of the columns. There are several evidence which show the resistance of Clostridiales (Gieg et al., 2008, Sherry et al., 2013, Mbadinga et al., 2011), Actinomycetales (Akbari and Ghoshal, 2015), Bacillales (Jayamani and Cupples, 2015, Engel and Gupta, 2014) to crude oil or similar hydrocarbonic compounds and their ability to bioremediation of these compound in aerobic or anaerobic conditions. Furthermore, the ratios of Xanthomonadales (*Stenotrophomonas acidaminiphila*), Pseudomonadales (*Pseudomonas* sp.), Burkholderiales (*Brachymonas petroleovorans*), Desulfobacterales (mainly *Desulfobacterium* sp.), Rhodospirillales (mentioned bellow) and Deferribacteriales (*Denitrovibrio acetiphilus* DSM 12809) were increased in the water layers of the columns treated by crude oil, which represent their high resistance to this compounds. *Stenotrophomonas acidaminiphila* is able to aerobic degradation of several aromatic hydrocarbons, such as dibenzothiophene and dihydroxynaphthalene (Papizadeh et al., 2011, Verma et al., 2011). Furthermore, Desulfobacterales, including the genera found in this study, are considered as one of most potent bacteria for degradation of hydrocarbons in anaerobic conditions (Sherry et al., 2013, Neria-González et al., 2006). Presence of Rhodospirillales, a group of phototrophic microorganisms with the ability to metabolise organic compounds (Garrity et al., 2005), in crude oil contaminated sites indicates their ability to use these compounds as their source of nutrients. Several Rhodospirillales were found in this study that their involvement in degradation of hydrocarbons have been reported in previous investigations; these genera included *Azospirillum* sp. (Muratova et al., 2005), *Dechlorospirillum* sp. (Sturchio et al., 2007), *Magnetospirillum* sp. (Widdel et al., 2010a), *Phaeospirillum* sp. (Liu et al., 2015), *Rhodocista sp.* (Brakstad and Bonaunet, 2006), *Thalassospira* sp. (Zhao et al., 2010). *Denitrovibrio acetiphilus* is a nitrate reducing bacterium found recently in oil reservoir columns and are potentially useful for degradation of hydrocarbons (Myhr and Torsvik, 2000, Zhang et al., 2012).

These changes in the microbial community means the microbial niches in all three layers of the columns were largely affected by the oil contamination. The metabolic pathway and

respiratory system of each order and species of microorganisms were obtained from authorised references (Whitman and Parte, 2009, Boone et al., 2011, Schink et al., 2011, Goodfellow et al., 2012). Based on these data, the microorganisms identified in each layer were classified into four metabolic categories: chemoheterotrophic bacteria, photoautotrophic organisms, photoheterotrophic bacteria and chemolithotrophic bacteria. Furthermore, the chemoheterotrophic bacteria were classified into aerobic bacteria, facultative anaerobic bacteria and obligate anaerobic bacteria based on the final electron acceptor through respiration. Since some species of bacteria do not show a same metabolic pathway as other members of an order, these species were categorized in another group based on its metabolic pathway. For instance, *Porphyrobacter tepidarius* is a photoautotrophic bacterium that belongs to Sphingomonadales while some other species of this order, including *Sphingopyxis terrae*, *Agrobacterium sanguineum*, *Sphingobium* sp. and *Sphingomonas* sp. shows heterotrophic activity.

Overall, the majority of the bacteria in all three layers of the control column belonged to chemoheterotrophic bacteria, and the highest ratio (78%) of these group of organisms were achieved at the bottom (Fig. 3.2.4). Although ratios of many of the chemoheterotrophic bacteria, such as Rhizobiales, Rhodobacterales, Planctomycetales, Sphingobacteriales, Cytophagales, Rubrobacterales, Spirochaetales, Thermoanaerobacterales, Thermotogales and two unclassified orders derived from Bacteroidetes and Opitutae dropped after oil contamination, the ratio of chemoheterotrophic microorganisms showed increases after these contaminations. These increases in the chemoheterotrophic microorganisms occurred as a result of increases in the ratios of Actinomycetales, Xanthomonadales, Deferribacteriales, Bacterioidales and Clostridiales in the water, Bacterioidales, Thermoanaerobacterales and Bacillales in the middle and Bacillales, Acholeplasmatales, lactobacillales, Thermoanaerobacterales, Thermotogales and Pseudomonadales in the bottom. The changes in the ratios of heterotrophic organisms were not severe in the water layers and middle layers of the columns treated by low concentrations of crude oil in comparison to the control. However, this value increased severely in highest concentration of crude oil used in this experiment (3%) and the highest change in this value was seen in the water layer of the column treated by 3% crude oil. Overall, these changes led to increases in the total abundances of chemoheterotrophic microorganisms in the columns treated by crude oil, and this value reached its peak in the column treated by 3% crude oil.

To more clarification, these chemoheterotrophic microorganisms were categorized based on the final electron acceptor molecule through catabolism of organic compounds into aerobic, facultative anaerobic and obligate anaerobic bacteria. Almost half portion (49%) of the chemoheterotrophic bacteria at the water layer of the control column belonged to aerobic bacteria, while this value decreased at the middle and bottom of the control column to 43% and 35%, respectively (Fig. 3.2.5). Although there were no big differences in the ratios of obligate anaerobic bacteria at these three layers, the ratios of facultative anaerobic bacteria increased from 4.8% at the water layer to 15.15 and 33.3% at the middle and bottom of the control column. Meanwhile, treatment by crude oil could change the ratios of aerobic and anaerobic heterotrophic organisms in all three layers of the columns. The treatments could led to increase in the ratios of aerobic bacteria and facultative anaerobic bacteria in the water layers of the columns and the highest ratio of aerobic bacteria was seen in the water layer of the column treated by 3% crude oil. There were no significant changes in the ratio of obligate anaerobic bacteria in the middle of the column, but there were small increases in the ratio of aerobic bacteria. However, the ratio of each aerobic, facultative anaerobic or obligate anaerobic bacteria was variable at the middle, and decreases in the ratios of several aerobic bacteria in the middle of the columns, such as Cytophagales, Rhizobiales and Rubrobacterales were compensated with increase in the ratios of some other aerobic bacteria such as Acholeplasmatales and Actinomycetales. Similarly, decreases in the ratios of facultative anaerobic bacteria, such as Sphingobacteriales and Verrucomicrobiales, and obligate anaerobic bacteria, such as Desulfobacterales, were compensated with growth in the ratios of other facultative and obligate anaerobic bacteria. These values were quite constant in the bottom except for some fluctuations in each value. Overall, these alterations in the ratios of aerobic and anaerobic bacteria at different layers of the columns treated by crude oil led to small decreases in the ratios of facultative anaerobic bacteria, replaced by aerobic bacteria.

Several groups of organisms, referred to as autotrophic organisms, are able to fix CO_2 as their main source of carbon. These organisms are categorized into several groups, including photoautotrophs, photoheterotrophs and chemolithotrophs based on the compound they use as the main source of hydrogen and electron to fix CO_2, such as H_2O, H_2S, H_2 and Fe. However, since the ratios of chemolithotrophic organisms in all layers were very small, we neglected

these microorganisms for further analysis. Based on the results, crude oil contamination decreased the total ratios of autotrophic organisms in each layer in comparison to the total abundances of microorganisms in that layer. The majority of the phototrophic organisms in the water layer of the control column belonged to oxygenic phototrophic bacteria (62.4), followed by green algae (34.7%) and photoheterotrophic bacteria (2.7%) (Fig. 3.2.6). Although eukaryotic algae become dominant at both middle and bottom of the control column (59,2% and 50%, respectively) and the level of oxygenic phototropic bacteria decreased to less than 40% at both deep layers, these oxygenic phototrophic organisms were still dominant in these layers. Treatment by crude oil, however, lead to decreases in the ratios of both oxygenic phototrophic organisms (eukaryotic algae and oxygenic phototropic bacteria) in all layers. The reduction was raised mainly because of decreases in the ratios of Chroococcales (only in the water layer), Sphingomonadales, Oscillatoriales, Nostocales, Naviculales (only in the bottom) and an unclassified diatom derived from Bacillariophyceae. Inversely, the treatment by crude oil provoked the proliferation of photoheterotrophic bacteria in all three layers. The ratio of these organisms increased in the water layer of the columns treated by crude oil, reaching its peak in the water layers treated by 3% crude oil. Although there were some fluctuations, the ratios of these organisms showed an additive trend in the columns treated by crude oil in comparison to the control. Growth in the ratio of an order of purple non sulfur bacteria (Rhodospirillales) in all three layers and an order of green non sulfur bacteria (Herpetosiphonales) in the middle of the columns were the main reason of increases in the ratio of photoheterotrophic bacteria in these columns. It appears that some constituents of crude oil were used as substrate for photoheterotrophic metabolism of these order of bacteria, and therefore exposure to these compounds could improve their growth ratios in these layers.

To determine the overall effects of crude oil contamination in these columns, the total microbial diversity and their ratios in each column were investigated without dividing the columns into layers. Since the DNA samples obtained from the water layer and the middle of the columns treated by 0.1% crude oil and the sample obtained from the bottom of the column treated by 1% crude oil were missed in this experiment, three concentrations which their data are available for all three layers were chosen for this analysis. This analysis showed that the treatments with crude oil could increase the total abundances of microorganisms in these columns, and increases in the concentrations of the treatment intensified the growth in this value (Table 3.2.2). These treatments had positive effects on the ratios of chemoheterotrophic bacteria in

these columns, reached its peak in the tubes treated by 3% crude oil (Fig. 3.2.2). The decreases in the ratios of photoautotrophic microorganisms was mainly because of decreases in the ratios of oxygenic phototrophic bacteria (Fig. 3.2.4). However, the ratios of photohetrotrophic bacteria reached from 0.9% in the control to 26.8% in the column treated by 3% crude oil (Fig. 3.2.4).

Overall, this study showed that crude oil pollution in a fresh water lake microcosm increases the total abundances of microorganisms in each layer and in whole column, mostly belongs to chemoheterotrophic bacteria. Furthermore, this contamination increases the ratio of aerobic bacteria mostly in the water layer of the microcosm. This conditions are along with decreases in the ratios of eukaryotic and prokaryotic oxygenic photoautotrophic organisms, which supply oxygen for the aerobic heterotrophic organisms. Furthermore, these phototrophic organisms are considered as the main primary food producers of the aquatic environments which supply the organic compounds required for higher heterotrophic living beings. This study also showed that increases in the volume of contamination make severe the effects on the ratio of microbial communities in the environment. These increases in the ratio of chemoheterotrophic organisms and decreases in the ratios of oxygenic phototrophic organisms can finally affect the activity of higher aquatic living beings, including fishes and aquatic invertebrates (Bouwman et al., 2013). However, since crude oil treatment increased the ratios of the chemoheterotrophic and photoheterotrophic bacteria, these changes can ultimately help the nature to reduce the hydrocarbon overloading in the area (Megharaj et al., 2011). It is worthwhile to mention that the type and ratio of hydrocarbon degrading bacteria are affected by the time of sampling, temperature, salinity, pH, availability of oxygen, bioavailability of nutritional elements (such as sulphur and nitrogen) and the extent of oil contamination (Atlas, 1981). Although all adverse aspects of oil pollutants on microbial diversity are not clearly known, it depends on the oil chemical composition and microbial species of the habitat (Sathishkumar et al., 2008). Furthermore, because the natural hydrocarbon degradation process relies on a complex interaction between the members of a community, more investigations are required to detect the role of each member in this community. The dynamic and continuous alternations of microbial communities in such complex environments can be tracked using molecular based methods.

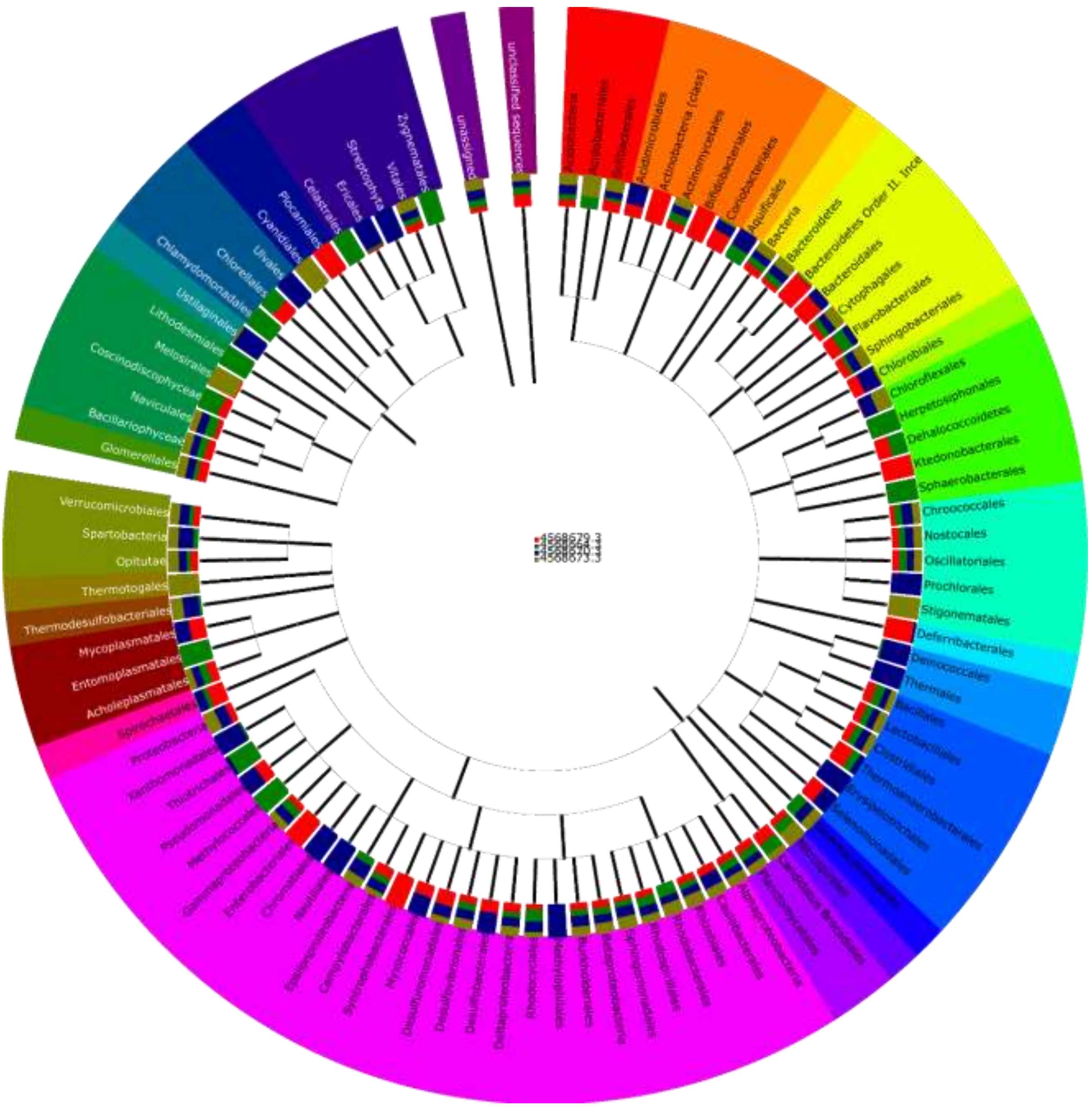


Figure 3.2.5 The microbial diversity in middle layer of winogradky column contaminated with different concentrations (0%, 0.1%, 1% and 3%) of crude oil. All microbial orders belonging to a phylum are indicated using different colours.

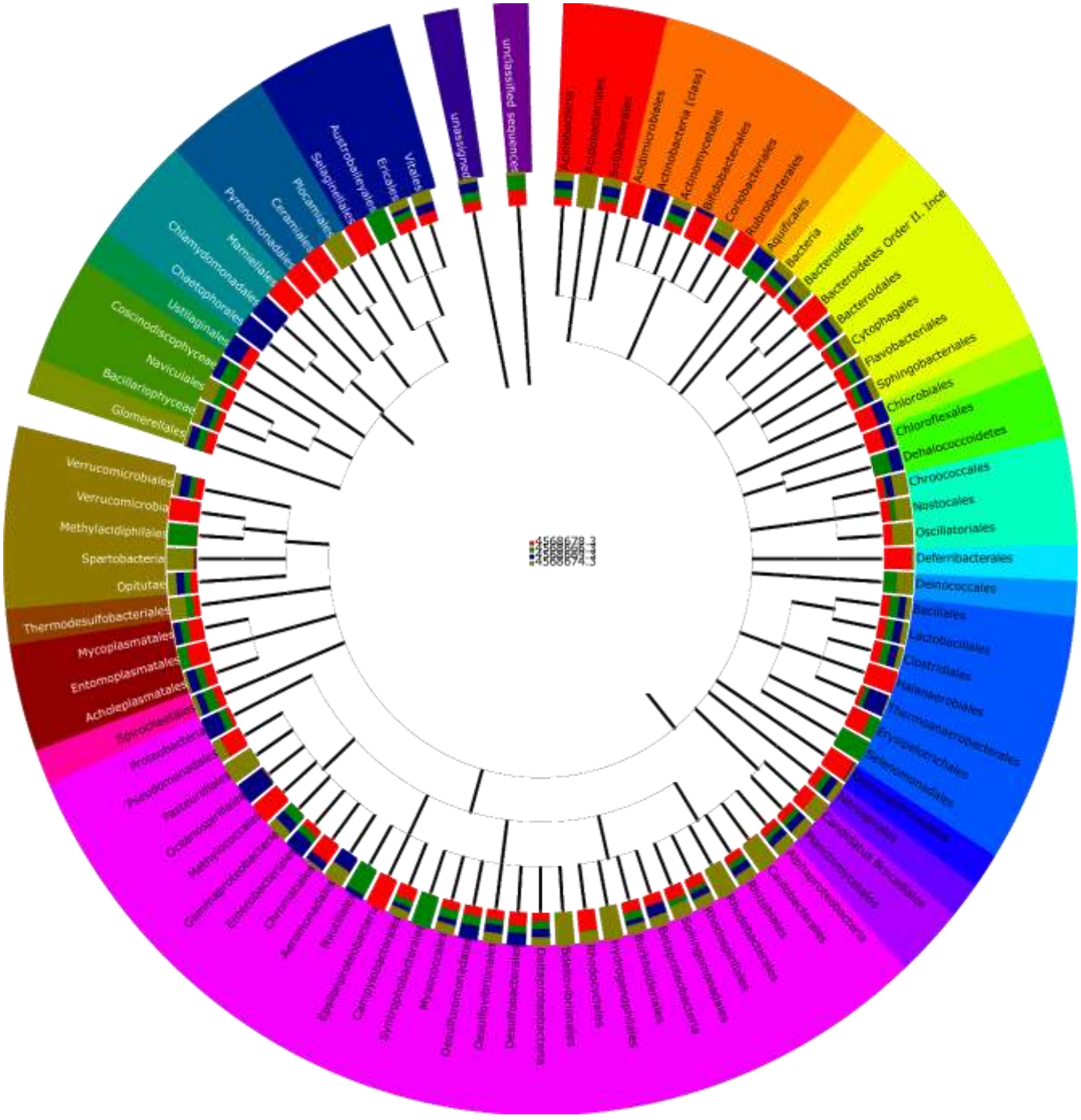


Figure 3.2.6 The microbial diversity in bottom of winogradky column contaminated with different concentrations (control, 0.1%, 0.5% and 3%) of crude oil. All microbial orders belonging to a phylum are indicated using different colours.

3.3 The biodiversity changes in the microbial population of soils contaminated with crude oil

Published in: Current Microbiology. 72: 663-670.

3.3.1 Abstract

Crude oil spills resulting from excavation, transportation and downstream processes can cause intensive damage to living organisms and result in changes in the microbial population of that environment. In this study, we used a pyrosequencing analysis to investigate changes in the microbial population of soils contaminated with crude oil. Crude oil contamination in soil resulted in the creation of a more homogenous population of microorganisms dominated by members of the Actinomycetales, Clostridiales and Bacillales (all belonging to gram positive bacteria) as well as Flavobacteriales, Pseudomonadales, Burkholderiales, Rhizobiales and Sphingomonadales (all belonging to gram negative bacteria). These changes in the biodiversity decreased the ratios of chemoheterotrophic bacteria at higher concentrations of crude oil contamination, with these being replaced by photoheterotrophic bacteria, manly Rhodospirillales. Several of the dominant microbial orders in the crude oil contaminated soils are able to degrade crude oil hydrocarbons and therefore are potentially useful for remediation of crude oil in contaminated sites.

3.3.2 Introduction

Crude oil is considered as one of the main sources of energy in the world, and growth in population and industries increases the demand for more crude oil extraction (Hussain and Gondal, 2008). Leakage of oil into soil around the oil wells, pipelines and pumping stations is a normal phenomenon in the process of extraction of crude oil (Hussain and Gondal, 2008). Furthermore, crude oil transportation and downstream processes are also high risk points for crude oil contamination (Nicolotti and Egli, 1998, Dorn et al., 1998). Although the level of soil contamination with crude oil depends on the site location, this value can rise to around 10% w/w (Gogoi et al., 2003). Contamination can be spread into other natural habitats by superficial waters or floods especially in the rainy seasons (Zakaria et al., 2002). Crude oils consist of

different aliphatic and aromatic hydrocarbons, which are rich in carbon and hydrogen but deficient in other nutritional elements (Sathishkumar et al., 2008, Morgan and Watkinson, 1994, Shi et al., 2010). Since these compounds can be potentially poisonous for different eukaryotic and prokaryotic cells, soil contamination by crude oil can lead to immediate changes in the microbial composition of the given environment (Abbasian et al., 2015a). Furthermore, because oil spillages change the chemical composition of soil and thereby the nutrient availability to microorganisms, contamination can lead to prolonged effects on the microbial diversity and their abundances in the contaminated sites (Megharaj et al., 2011). These changes in microbial population depend heavily on the composition of the total microbial community present in the site, the chemical composition of the crude oil and the physicochemical factors governing the particular environment (Rahman et al., 2002).

Since microorganisms are the most effective factors in the natural degradation of crude oil and industrial bioremediation, understanding the changes in the microbial diversity resulting from contamination of soil with crude oil can be useful for the selection of the most effective hydrocarbon degrading microorganisms for crude oil remediation. The most abundant microorganisms in a soil sample contaminated by crude oil are potentially beneficial in facilitating the removal of crude oil contamination (Balba et al., 1998b, Odokuma and Dickson, 2004, Gallego et al., 2001). In this study, the changes in microbial diversity and their abundances were analysed using a high throughput pyrosequencing strategy in order to understand these dominant microorganisms in the soils contaminated with crude oil.

3.3.3 Material and methods

3.3.3.1 Preparation of the column and oil spiking

A Gauthier was used to collect the superficial soil samples (6.8 pH at room temperature, 2.8% moisture and 612 µS conductivity) in order to save the integrity of soil samples. The soil samples were carefully transferred to suitable columns matched to the weight of soil samples (Fig. 1). Among the five columns, one column was used as control without any treatment, and four others were treated with different concentration (0.5%, 2.5%, 5% and 10%) of crude oil. Overall, two sets of these columns were prepared, allowing the reproducibility of the results to be determined; these columns were incubated in room temperature (25 ^{o}C) for one week.

3.3.3.2 DNA extraction and Pyrosequencing process

Following the incubation period, several samples were obtained from different parts of each column. The samples obtained from each column were mixed thoroughly to make a homogenous sample. The DNA from the soils was extracted using a power soil DNA kit (MO BIO), according to manufacturer's instructions. First, after a gentle vortex of 0.25 g sediment samples in tubes containing beads and bead buffer, the contents were mixed with 60 µl solution C1. Following a vigorous vortexing for 10 minutes using a bead beating machine and centrifuging at 10,000 x g for 30 S, the supernatants (400-500 µl) were processed according to the manufacturer's instructions. The levels of DNAs obtained were quantified by a quantifluor dsDNA system (Promega) and were adjusted to a minimum concentration of 100 ng/µl.

The genomic DNAs were sequenced on the GS-FLX platform located at the Australian Genome Research Facility (AGRF). First, the DNAs were amplified using the primers directed for 16SrRNA of bacteria (referred to as 16S: 27F - 519R; 5′-AGAGTTTGATCMTGGCTCAG-3′ and 5′-GWATTACCGCGGCKGCTG-3′) and archaea (referred to as 16S: 341F-806R; 5′-CCTAYGGGRBGCASCAG-3′ and 5′-GGACTACNNGGGTATCTAAT3-′) as well as 18SrRNA of fungi (referred to as ITS: 1F-2R; 5′-CTTGGTCATTTAGAGGAAGTAA-3′ and 5′-TCCTCCGCTTATTGATATGC-3′) and algae (referred to as Euk3191F-EukBR; 5′-CTGGTTGATCCTGCCAG-3′ and 5′-ACCAGACTTGCCCTCC-3′). After a DNA quantification of the PCR products by fluorometry and qPCR, the DNA fragments were sequenced using the GS-FLX platform by the Australian Genome Research Facility (AGRF).

3.3.3.3 Data Analysis

The raw data files obtained from pyrosequencing were uploaded into the MG-RAST (Metagenome Rapid Annotation using Subsystem Technology) (Glass and Meyer, 2011, Abbasian et al., 2015b), and the microbial diversity and their abundances were analysed using the Ribosomal Data Project II (RDP II) annotation source. To minimize the mistakes through analysis on the microbial diversity using 16SrRNA, the data were filtered by minimum

percentage identity cut off (97%), E-value cutoff (1e-5) and minimum alignment length cutoff (50 bp) (Abbasian et al., 2015d).

The microbial diversity and their abundances were analysed further using excel software to analyse the Shannon's index and to investigate the ecological changes as a result of crude oil treatment. The Shannon index (a minus sum of the proportion (n/N) of the number of one particular species in a community divided by the total number of individuals in that community multiplied by the natural logarithm of this proportion) reflects the number of species in a dataset and their distribution in a community (Keylock, 2005). To increase the accuracy of our analysis, the metabolic (chemoheterotrophic, photoautotrophic, photoheterotrophic and chemolithotrophic) activity of all microbial strains identified were extracted from authorised references (Whitman and Parte, 2009, Boone et al., 2011, Schink et al., 2011, Goodfellow et al., 2012), and the strains with a same metabolic activity and same phylogenetic order were grouped together.

3.3.3.4 Nucleotide sequence accession numbers

These metagenomic data are publicly available in the MG-RAST system as F44 (control), F45 (treated with 2.5% crude oil), F47 (treated with 5% crude oil), F48 (treated with 0.5% crude oil) and F49 (treated with 10% crude oil) under project identifiers 4576644.3., 4576645.3., 4576646.3., 4576647.3. and 4576648.3., respectively.

3.3.4 Results

The numbers of reads after quality control, mean length and the numbers of reads fitted in an operational taxonomic unit (OTU) for all samples are listed in table 3.3.1. A full list of the microbial orders found in these columns are given in figure 3.3.1. According to the saturated plateaus in the rarefaction curves obtained from the analysis of the control soil sample and the samples treated with different concentrations of crude oil, the abundances of microbial population in these sampling are representative of those present in the columns (data are not shown). Moreover, according to the principal coordinate analysis (PCoA) (data are not shown) these treatments considerably altered the biodiversity present in the treated columns in comparison to the control sample. In total, the abundances of microorganisms in the sediments

treated with 0.5% and 2.5% crude oil were increased in comparison to the control. Although this value reached its peak in the soil sample treated by 0.5% crude oil, there were gradual decreases in the total abundances of microorganisms in higher concentrations (2.5% and 5%) to a level slightly more than the control). However, there was a small rise again in the abundance of microorganisms when the soil sample were treated with 10% crude oil. In addition to total abundances of microorganisms, the treatments decreased the numbers of microbial orders in these samples from 59 in the control (0%) to 46 in the soil contaminated with 10% crude oil. Furthermore, there were changes in the ratios of microbial orders in the samples when compared with the total abundances of microorganisms in the corresponding sample. These changes in the abundances of microbial orders were significant after analysis with Shannon index (Table 3.3. 2).

Sample	Number of reads after quality control	Mean length	Number of reads fitted in an operational taxonomic unit (OTU)
F44 (4576644.3.)	2166	428 ± 144 bp	842
F45 (4576645.3.)	3871	450 ± 119 bp	1513
F47 (4576646.3.)	3331	449 ± 115 bp	1342
F48 (4576647.3)	2276	457 ± 128 bp	929
F49 (4576648.3)	3070	448 ± 118 bp	979

Table 3.3.1. The statistical data obtained from the pyrosequencing analysis of the control soil and the soils treated with different concentrations (0.1%, 0.5%, 2.5%, 5% and 10%) of crude oil.

	0.0%	0.5%	2.5%	5.0%	10.0%
H	2.71	2.83	2.70	2.34	2.45
D	14.98	17.00	14.90	10.37	11.56

Table 3.3.2. Shannon index (H) and effective numbers (D) of orders in the control soil and the soils treated with different concentrations (0.1%, 0.5%, 2.5%, 5% and 10%) of crude oil.

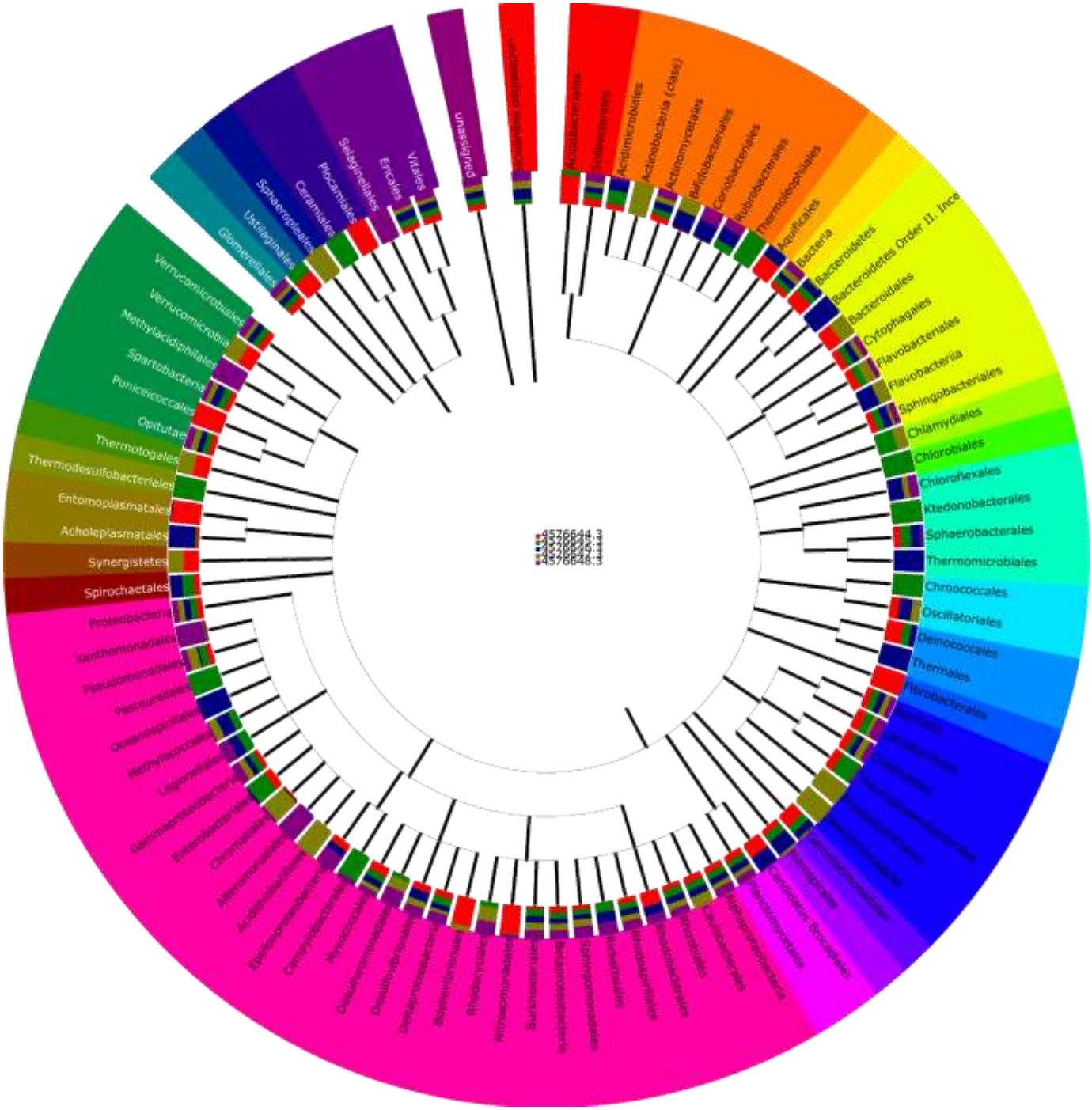


Figure 3.3.1. The microbial diversity (at the order level) in the control column (4576644.3) and the columns treated by different concentrations (0.5%, 2.5%, 5% and 10%) of crude oil (4576645.3, 4576646.3, 4576647.3, 4576648.3, respectively).

Among these microorganisms, Actinomycetales (20%) were the dominant microorganism in the control soil, followed by Cytophagales (7.4%), Solibacterales (3.6%), Glomerellales (3.5%), Planctomycetales (3.4%), Verrucomicrobiales (3.2%), Clostridiales (3%), Bacillales (2.5%) and many other orders with less than 2.5% abundance (Fig. 3.3.2). However, the ratios of microorganism changed in the samples treated with different concentrations of crude oil. In the soil sample treated with 0.5% crude oil (Fig. 3.3.2) these were, Actinomycetales (25%), Flavobacteriales (6.1%), Glomerellales (2.9%), Planctomycetales (2.8%), Verrucomicrobiales (2.7%), Rhizobiales (2.6%), Rhodospirillaes (2.5%), Cytophagales (2.5%), Bacillales (2.1%) and many other orders with less than 2% of the population (Fig. 3.3.2).

The ratio of Actinomycetales (37.21%) reached the highest amount in the soil spiked with 2.5% crude oil, followed by Rhizobiales (3.4%), Rhodospirillales (3%), Flavobacteriales (2.9%), Planctomycetales (2.9%), Cytophagales (2.5%), Clostridiales (2.3), Bacillales (2%) and many other orders that were less than 2% of the population (Fig. 3.3.2). These value changed in the soils contaminated with higher levels of crude oil, as well. Although the ratio of Actinomycetales decreased in the soil sample treated with 5% crude oil, still it was the dominant microorganism in this soil, followed by Sphingomonadales (5.1%), Rhizobiales (4.6%), Burkholderiales (3.4%), Rhodospirillales (3.0%), Flavobacteriales (2.9%), Clostridiales (2.3%), Rhodocyclales (2.2%), Bacillales (2.1%) and many other orders less than 2% abundance. Furthermore, these changes in the diversity and ratios of microorganisms in the soil sample treated by 10% crude oil were much severe than other samples. In this sample Sphingomonadales (16%) was nominated as the dominant microorganism and Actinomycetales (10.2%) was placed in the second step, followed by Rhodospirillales (6.3%), Rhizobiales (2.8%), Burkholderia (2.4%), Planctomycetales (2%) and many other orders with less than 2% abundances.

Based on the literature and microbial classification handbooks (Boone et al., 2011, Goodfellow et al., 2012, Schink et al., 2011, Whitman and Parte, 2009), all microbial strains identified in this study were classified into four physiological categories (chemoheterotrophs, photoautotrophs, photoheterotrophs and chemolithotrophs) to analyse the ecological relationships between microorganisms in the soil samples. Thereafter, these strains were re-classified into their orders to facilitate the statistical analysis. Since the members of microbial

orders show sometimes very diverse physiology, this procedure of microbial classification from strains to orders reduces the errors in calculations. Based on this study, the majority of microorganisms in the control and all soils treated with different concentrations of crude oil belonged to chemoheterotrophic bacteria (Fig. 3.3.3). However, the soil treated with 10% crude oil showed a significant decrease in the ratio of chemoheterotrophic bacteria (83%), replaced by photoheterotrophic bacteria (17%).

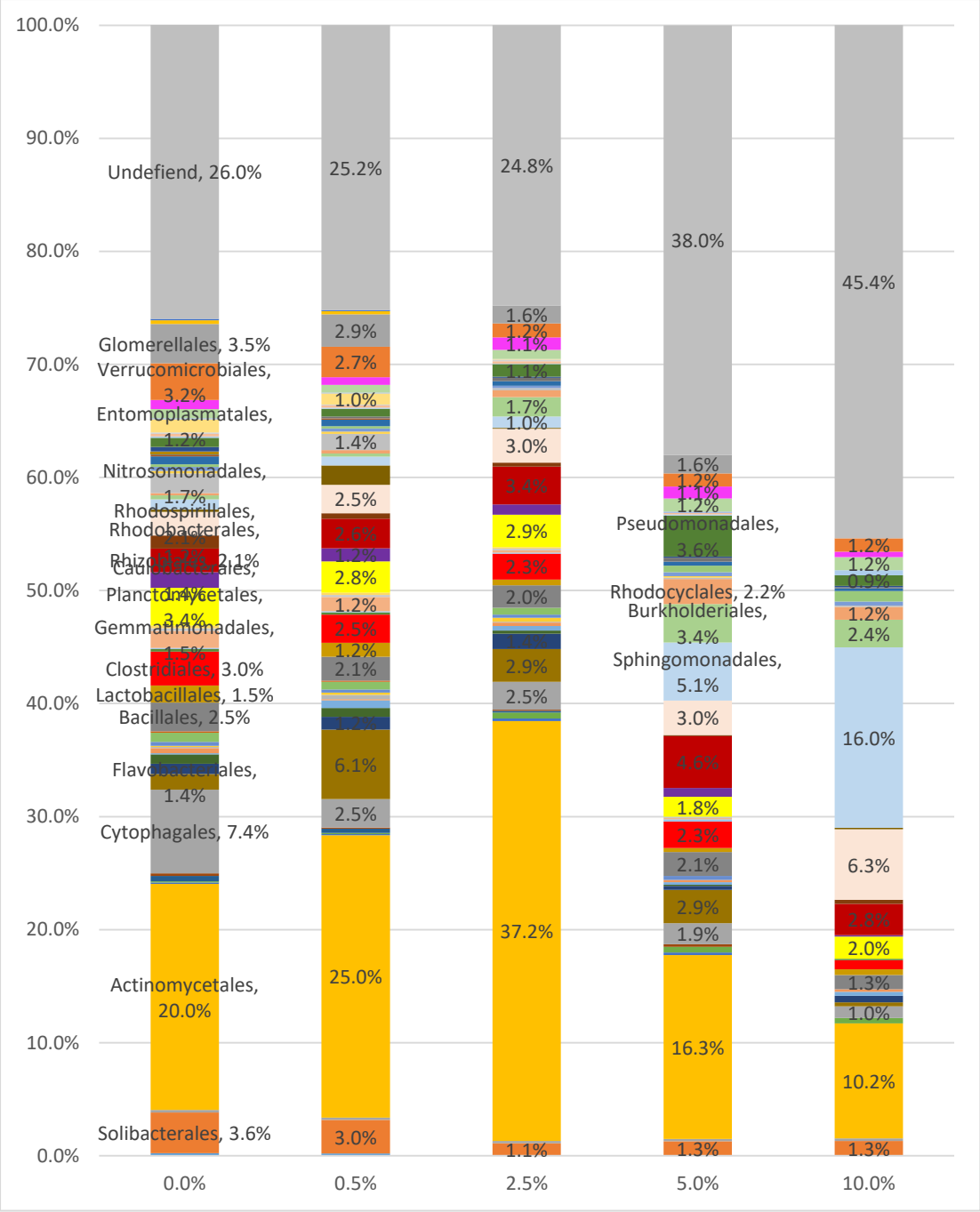


Figure 3.3.2. Ratios of microbial orders in the soils contaminated with different concentrations (0%, 0.5%, 2.5%, 5.0% and 10%) of crude oil.

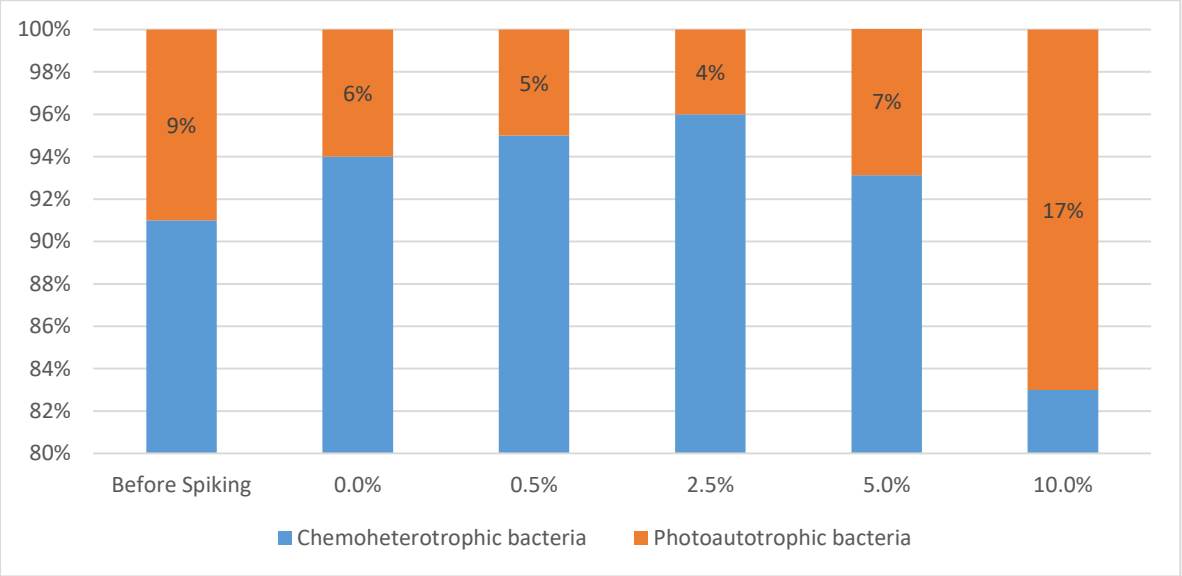


Figure 3.3.3. Ratios of chemoheterotrophic and phototrophic microorganisms present in the soils contaminated with different concentrations shown as a percentage of the total population (0%, 0.5%, 2.5%, 5.0% and 10%) of crude oil.

3.3.5 Discussion

In this study, a pyrosequencing study was employed to investigate changes in the microbial diversity and the ratios of their populations in soils contaminated with crude oil. Based on the results, crude oil contamination showed severe effects on the total microbial population and their biodiversity in soils, and these effects depended heavily on the concentration of the crude oil. Although the total abundance of microorganisms in the sample treated with 0.5% crude oil increased in comparison to the control sample, further increases in the levels of crude oil contamination led to decreases in the total abundances of microorganism. Furthermore, these treatments affect negatively the numbers of microbial orders identified in these soil subjects. Therefore, crude oil contamination in soil created a more homogenous environment in terms of microbial diversity, and many of the microorganisms susceptible to constituents of crude oil were lost from these environment. Several microbial orders, such as Fibrobacterales, Thermoanaerobacterales, Nautiliales, Enterobacteriales, Thermotogales, Ustilaginales and Plocamiales, showed a high susceptibility to this treatment and vanished from the soils spiked

with low concentrations of crude oil (0.5% and 2.5%). This list was extended to Aquificales, Thermomicrobiales, Chroococcales (a Cyanoobacter), Deinococcales, Candidatus Brocadiales, Acholeplasmatales and Entomoplasmatales (in the soil samples treated with 5% crude oil) and to Bacteroidales, Oscillatoriales (an algae), Selenomonadales, Gemmatimonadales, Nitrospirales and Glomerellales (in the soil samples treated with 10% crude oil).

Furthermore, except for some orders, the ratios of the majority of microorganisms decreased in the soil spiked with crude oil in comparison to the control samples. These changes in microbial diversity led to transformations in the abundant microorganisms in these soils. Members of Actinomycetales, mostly *Mycobacterium* sp., Nocardia sp. and Pseudonocardia sp.), were the dominant microorganisms in the control soil and its ratio even increased in the soils treated with 0.5% and 2.5% crude oil. Although the ratio of Actinomycetales decreased in higher levels of treatment, the members of this order were still one of the most dominant microorganisms in the soils treated with crude oil. A high level of resistance to crude oil is very common among members of this order, and many genera in this order are commonly used for the degradation of different types of hydrocarbons (Kweon et al., 2011, Kim et al., 2015, Zeinali et al., 2008). Members of Clostridiales (*Clostridium* sp., *Symbiobacterium* sp. and *Oxobacter* sp.) and Bacillales (*Bacillus* sp. and *Brevibacillus* sp.) were two other gram positive orders with the ability to resist high levels of crude oil contamination, while their ratios decreased significantly in the soils treated with the10% treatment. While members of Clostridiales are involved in the degradation of alkanes in anaerobic conditions (Gieg et al., 2008, Sherry et al., 2013, Mbadinga et al., 2011), genera of Bacillales degrade these compounds in aerobic conditions (Margesin and Schinner, 2001b, Ghazali et al., 2004, Cubitto et al., 2004).

Several groups of gram negative bacteria, including Flavobacteriales, Pseudomonadales, Burkholderiales, Rhizobiales and Sphingomonadales, were identified as dominant microorganisms in these soil samples as well. The highest ratio of Flavobacteriales, mainly *Flavobacterium* sp., Pseudomonadales (mainly *Pseudomonas* sp.), Burkholderiales (mainly *Variovorax* sp. and *Brachymonas petroleovorans*) and Rhizobiales (mainly *Bradyrhizobium* sp.) were observed in the soils treated by 0.5% crude oil, but higher levels of treatment decreased their values. Among these group of aerobic bacteria, members of Pseudomonadales (Pacwa-Płociniczak et al., 2014, Matsui et al., 2014), Burkholderiales (Dobslaw and Engesser,

2014) and Flavobacteriales (Vinas et al., 2005, Balba et al., 1998a, Gentili et al., 2006, Kaplan and Kitts, 2004) have been commonly used for the bioremediation of crude oil contaminated subjects.

Inversely, increases in the levels of crude oil treatment raised the ratio of Sphingomonadales (mainly *Sphingomonas* sp.) in the soils, and members of this order became the most dominant microorganisms in the soil treated with 10% crude oil. This increase in the ratio of *Sphingomonas* sp. indicates its ability to resist high crude oil pollution and its ability to degrade different types of hydrocarbons as reported in previous studies (Wang et al., 2012b, Dai and Copley, 2004, Colombo et al., 2011, Liang and Lloyd-Jones, 2010, Koukkou and Vandera, 2011). Rhodospirillales showed a similar distribution, and members of this order became the second most dominant microorganisms in the soils treated with 10% crude oil. Rhodospirillales show a photoheterotrophic metabolism in which electrons and hydrogen required for CO_2 fixation are originate from some organic molecules (Garrity et al., 2005). Two orders of Rhodospirillales (*Azospirillum* sp. and *Phaeospirillum* sp.) were found in these samples, and their involvement in the degradation of crude oil has been shown in previous studies (Muratova et al., 2005, Liu et al., 2015).

These changes in the diversity and ratios of microorganisms affects the total microbial physiology of the soil samples. Although chemoheterotrophic bacteria were dominant in both control soil and soils spiked with crude oil, this value decreased with higher concentrations of crude oil contamination, where they were replaced by photoheterotrophic bacteria, manly Rhodospirillales. Furthermore, these treatments did not show significant effects on the ratios of both oxygenic and non-oxygenic phototrophic microorganisms. Overall, this study indicated that several types of soil microorganisms are highly susceptible to crude oil contamination. However, there were several other microbial orders, especially Sphingomonadales, which are known as crude oil hydrocarbon degrading organisms that are able to resist these conditions. In addition to adaptations for resistance to the toxic effects of different classes of hydrocarbons and their resistance to heavy metals, which are naturally found in crude oil, these microorganisms are able to produce the enzymes required for the degradation of aliphatic and aromatic hydrocarbons. Such microorganisms can be used as potential candidates for remediation of crude oil in contaminated sites. Furthermore, investigation on the genetic

capacity of these microorganisms can provide significant information regarding the genes involved in hydrocarbon bioremediation as well as the production of bio-surfactants.

3.4 Multiwall carbon nanotubes increase the microbial community in crude oil contaminated fresh water sediments

Published in: Science of the Total Environment, 2016 Jan 1;539:370-80. doi: 10.1016/j.scitotenv.2015.09.031. Epub 2015 Sep 12.

3.4.1 Abstract

Since crude oil contamination is one of the biggest environmental concerns, its removal from contaminated sites is of interest for both researchers and industries. In situ bioremediation is a promising technique for decreasing or even eliminating crude oil and hydrocarbon contamination. However, since these compounds are potentially toxic for many microorganisms, high loads of contamination can inhibit the microbial community and therefore reduce the removal rate. Therefore, any strategy with the ability to increase the microbial population in such circumstances can be of promise in improving the remediation process. In this study, multiwall carbon nanotubes were employed to support microbial growth in sediments contaminated with crude oil. Following spiking of fresh water sediments with different concentrations of crude oil alone and in a mixture with carbon nanotubes for 30 days, the microbial profiles in these sediments were obtained using FLX-pyrosequencing. Next, the ratios of each member of the microbial population in these sediments were compared with those values in the untreated control sediment. This study showed that combination of crude oil and carbon nanotubes can increase the diversity of the total microbial population. Furthermore, these treatments could increase the ratios of several microorganisms that are known to be effective in the degradation of hydrocarbons.

3.4.2 Introduction

Annually millions of tons of crude oil are extracted around the world, and demand is increasing as a result of growth in population and industries (Hussain and Gondal, 2008). Any crude oil processing, from extraction to consumption, is a potential source of environmental contamination (Hussain and Gondal, 2008, Abbasian et al., 2015a). Crude oils consist mainly (80-90%) of aliphatic and aromatic compounds and partly of more complex compounds such as resins and asphaltenes (Sathishkumar et al., 2008, Morgan and Watkinson, 1994, Shi et al.,

2010). Environmental contamination by crude oil therefore, poses a serious risk to the inhabitants of that environment, especially microorganisms, due to the toxicity of some of its constituents (Morgan and Watkinson, 1994, Sánchez-Palencia González, 2011). The environmental contamination due to crude oil increases the ratio of carbon to nitrogen and phosphorus (C:N:P ratio) thereby creating an imbalance of nutrients making it unfavourable for the growth of microorganisms (Abbasian et al., 2015a). Furthermore, some constituents of crude oil and/or their breakdown products, especially monocyclic and polycyclic aromatic hydrocarbons are toxic to cells probably as a result of damage to the cell membranes and mutagenic effects on DNA (Reineke et al., 2006, Sikkema et al., 1995). Therefore, US EPA (US Environmental Protection Agency) has included these components as priority contaminants of natural resources (Yan et al., 2004).

By now, several mechanical, chemical and biological approaches have been designed for the removal of oil contaminations from a given environment (Fingas, 2010, Fingas, 2012). Bioremediation, is a natural process in which several living organisms, especially microorganisms, degrade oil components in a safe, cost effective and ecofriendly manner (Megharaj et al., 2011). Since the rate of hydrocarbon degradation in this system is mainly dependent on type and ratio of hydrocarbons in crude oil and several other physicochemical conditions, such as availability of oxygen, and essential nutrients, salinity, pH and temperature (Sierra-García et al., 2014, Sánchez-Palencia González, 2011), employment of a special microorganism or a consortium of microorganisms adapted to the given environment is critical in order to facilitate a high removal rate. Although many microorganisms are highly sensitive to oil hydrocarbons (Tang et al., 2011, Hubálek et al., 2007), some microorganisms are adapted to utilise these compounds through the production of degrading enzymes and can alter the composition of their cell membrane in response to exposure to hydrocarbons (McCammick et al., 2010). These bacteria increase the amount of trans- fatty acids in their cell membranes to improve their resistance to hydrocarbons (McCammick et al., 2010). As a result of their resistance to these contaminants, these organisms become dominant in a habitat polluted by crude oil or any other hydrocarbons (Nolvak et al., 2012, Boopathy, 2000). Some of these microorganisms are potentially useful for removal of hydrocarbons through bioremediation process. However, since no microbial species or microbial consortium has been introduced with the ability to grow in many different physicochemical conditions, growth and activity of indigenous hydrocarbon degrading microorganism can be the best way to increase the oil

bioremediation in an environment (Abbasian et al., 2015a). In addition to the hydrocarbon degrading microorganisms, some other microorganisms are able to use the by-products produced by the degradation of hydrocarbons (Abbasian et al., 2015a). Therefore, application of techniques to increase the microbial populations in crude oil contaminated sites may be useful for enhancing the bioremediation process.

The insolubility of hydrocarbons is one of the biggest challenges which microorganisms have to cope with for their degradation (Morgan and Watkinson, 1994, Sánchez-Palencia González, 2011). However, since MWCNMs display a strong hydrophobicity and great sorption capacity due to their large reactive surface areas, they can be used as sorbents of different hydrophobic organic compounds, including crude oil molecules (Mauter and Elimelech, 2008). Although there are several reports of cell toxicity of nanotubes on both prokaryotic (Arias and Yang, 2009, Cortes et al., 2014, Zardini et al., 2013, Dong et al., 2013) and eukaryotic (Wang et al., 2013a, Madani et al., 2013, Scott-Fordsmand et al., 2008) cells due to interruption of cellular metabolic activities, production of oxidative stress and cell membrane damage (Wang et al., 2013a, Madani et al., 2013, Kang et al., 2008), in our previous study (unpublished data) we observed that low concentrations (0.1% and 0.5%) of multiwall carbon nanotubes (MWCNTs) can increase the microbial biomass in freshwater sediments. Therefore, we hypothesised that the addition of MWCNTs to the crude oil contaminated areas may increase the accessibility of hydrocarbons to microorganisms and enhance microbial growth and biomass. Since there is deficiency of nutrient sources in crude oil and the contaminated area, mostly the hydrocarbon degrading microorganisms and the microorganisms which use the metabolites of the hydrocarbon degrading bacteria are able to survive in these environments. Therefore, this strategy may lead to improve the crude oil bioremediation rate in the contaminated areas. In this study, we used a pyrosequencing approach to investigate the effects of combinations of different concentrations of crude oil and multiwall carbon nanotubes on microbial diversity and their abundances in spiked fresh water sediments.

3.4.3 Material and methods

3.4.3.1 Experimental setup

The sediment samples (collected from Mawson lakes; South Australia), were sieved through a 2 mm sieve to remove any macroscopic animals, plants and rubble. Twenty grams of the sediments (6.9 pH, 2.7% moisture and 615 µS conductivity) placed in sterile glass tubes (50 ml vol) were spiked with different concentrations of crude oil alone or in combination with multiwall carbon nanotubes (Table 1). Following spiking, the sediments were shaken for 10 days under diffused natural light at room temperature (21 ± 2 ^{o}C). Experiment was conducted in triplicate.

3.4.3.2 DNA extraction and pyrosequencing process

After the incubation period, sediments in each vial were blended completely for homogenisation, and 5 grams of homogenous sediments from each treatment and control were used for DNA extraction. DNA extraction was performed by the power soil DNA kit (MO BIO) according to manufacturer's instruction. The DNA samples were quantified by the quantifluor dsDNA system (Promega) and their concentrations adjusted to 100 ng/µl. After DNA amplification process using the primers targeted for bacterial 16S rRNA (referred to as 16S: 27F - 519R; AGAGTTTGATCMTGGCTCAG and GWATTACCGCGGCKGCTG), archaeal 16S rRNA (referred to as 16S: 341F-806R; CCTAYGGGRBGCASCAG and GGACTACNNGGGTATCTAAT), fungal 18SrRNA (referred to as ITS: 1F-2R; CTTGGTCATTTAGAGGAAGTAA and TCCTCCGCTTATTGATATGC) and algal 18S rRNA (referred to as Euk3191F-EukBR; CTGGTTGATCCTGCCAG and ACCAGACTTGCCCTCC), the DNA samples were (pyro)sequenced on the GS-FLX platform at the Australian Genome Research Facility (AGRF).

3.4.3.3 Data Analysis

The raw data obtained from GS-FLX platform sequencing were trimmed using QIIME (Quantitative Insight Into Microbial Ecology) (Kreisinger, 2014) to remove any duplicated or/and uninformative sequences (read length). The data were submitted in Metagenome Rapid Annotation using Subsystem Technology (MG-RAST) (Glass and Meyer, 2011), and were analysed using the Ribosomal Data Project II (RDP II) annotation source and with regards to minimum percentage identity cut off of 97%, E-value cutoff of $1e^{-5}$ and minimum alignment length cutoff of 50 bp. MG-RAST is able to perform statistical analysis for calculation of the

principal coordinate analysis (PCoA) and the rarefaction curves. Additional statistical analysis for calculation of Sharon's index and also changes in the physiologic and ecologic conditions were conducted in excel software. Several authorised references (Whitman and Parte, 2009, Boone et al., 2011, Schink et al., 2011, Goodfellow et al., 2012) were used to classify microbial orders into different physiological groups based on the energy and carbon sources (chemoheterotrophic bacteria, oxygenic photoautotrophic organisms, anoxygenic phototrophic bacteria, photoheterotrophic bacteria and chemolithotrophic bacteria) and on the final electron recipient factor (aerobic bacteria, facultative anaerobic bacteria and obligatory anaerobic bacteria). Since some microbial species show a different behaviour from other members of the corresponding order, the physiology of each species was analysed separately and in case of any difference, they were classified in a separate group for more accuracy in the analysis.

3.4.3.4 Nucleotide sequence accession numbers

The reads are publicly available in the MG-RAST system as F24, F25, F26, F27, F29, F30, F31, F32, F34, F35, F36, F37, F38, F40, F41 and F42 under project identifiers 4576625.3, 4576626.3, 4576627.3, 4576628.3, 4576630.3, 4576631.3, 4576632.3, 4576633.3, 4576635.3, 4576636.3, 4576637.3, 4576638.3, 4576639.3., 4576640.3., 4576641.3. and 4576642.3., respectively.

3.4.4 Results

After 10 days incubation, DNAs were extracted from the sediment samples and subjected to pyrosequencing. The numbers of raw reads and the reads after trimming and other statistical data of the pyrosequencing analysis are mentioned in Table 3.4.1. Following classification of the sediments into four groups based on the concentrations of carbon nanotubes (Table 3.4.2), the changes in abundances in each sample after treatment with different concentrations of crude oil were compared with the control sediment sample.

The saturated and end-point flatted plateau in the rarefaction curves showed that the numbers of each OTU obtained from the DNA sampling were reasonable enough to efficiently show the microbial diversity of that sediment samples (data not shown). Furthermore, the data obtained from bioinformatic analysis showed massive changes after treatment of sediments with crude

oil. The principal coordinate analysis (PCoA) showed that the changes in each treatment were large enough to alter their microbial diversity (data not shown). Furthermore, the results of calculation of Sharon's index (Table 3.4.2) showed also that the alpha diversity and therefore the effective numbers of microbial orders in the several sediments were different from that value in control sample (Table 3.4.2).

Sample	Number of reads after quality control	Mean length	Fitted in an operational taxonomic unit (OTU)
F40 (4576640.3.)	3,265	439 ± 127 bp	1,104
F41 (4576641.3.)	4,211	439 ± 127 bp	1,234
F42 (4576642.3.)	1,312	456 ± 125 bp	477
F24 (4576625.3)	3042	435 ± 115 bp	980
F25 (4576626.3)	2081	421 ± 125 bp	730
F26 (4576627.3)	3235	436 ± 121 bp	1000
F27 (4576628.3)	3952	438 ± 126 bp	1169
F29 (4576630.3)	3036	433 ± 137 bp	1007
F30 (4576631.3)	1934	432 ± 148 bp	628
F31 (4576632.3)	5852	434 ± 126 bp	1621
F32 (4576633.3)	661	441 ± 146 bp	298
F34 (4576635.3)	1900	435 ± 129 bp	650
F35 (4576636.3)	2357	442 ± 119 bp	671
F36 (4576637.3)	1815	435 ± 135 bp	563
F37 (4576638.3)	1479	436 ± 132 bp	488

Table 3.4.1. The statistical data obtained from the pyrosequencing analysis of the sediments contaminated with mixture of crude oil and multiwall carbon

These treatments could change the total abundances of reads in the sediments depending on the concentrations of crude oil and nanotubes (Fig. 3.4.1). Overall, the abundances of reads increased in all sediments treated with different concentrations of crude oil and carbon nanotubes. The maximum numbers of reads in each group belonged to CNTs 0%-Crude oil 2.5% for group a, CNTs 0.1%-Crude oil 1% for group b, CNTs 0.5%-Crude oil 2.5% for group c and CNTs 1%-Crude oil 5% for group d. Furthermore, the maximum changes in the effective orders in group a, b, c and d in comparison to the control were obtained from the sediments

treated with CNTs 0%-Crude oil2.5%, CNTs 0.1%-Crude oil5%, CNTs 0.5%-Crude oil1% and CNTs 1%-Crude oil2.5% (Table 3.4.2). The sediments with higher values of reads showed changes in the numbers of effective orders in comparison to the control. Therefore, we used the sediments with maximum numbers of reads in each group for further studies.

Index	**Group a**				**Group b**				**Group c**				**Group d**			
	CNTs0%				CNTs 0.1%				CNTs 0.5%				CNTs 1%			
	Oil0%	Oil1%	Oil2.5%	Oil5%	Oil0.5%	Oil1%	Oil2.5%	Oil5%	Oil5%	Oil1%	Oil2.5%	Oil5%	Oil0.5%	Oil1%	Oil2.5%	Oil5%
H=	2.73	2.75	2.57	2.61	2.93	2.62	2.83	3.00	2.89	3.05	3.01	2.99	3.01	2.97	3.07	2.99
D=	15.38	15.57	13.10	13.64	18.72	13.69	17.02	20.15	18.00	21.09	20.36	19.98	20.35	19.59	21.53	19.90

Table 3.4.2. Sharon's index (H) and effective numbers (D) of orders in the control soil and the soils treated with different concentrations of carbon nanotubes and crude oil.

In addition to the effects on the total abundances of microorganism in the treated sediment, the combination of different concentrations of crude oil and CNTs could affect the ratios of each microbial orders in each sediment compared to the control sediment. The ratio of each microbial order in comparison to the total abundances of microorganisms in the corresponding sediment are shown in figures 3.4.2, 3.4.3, 3.4.4 and 3.4.5. A large percentage (29.2%- 41.1%) of microorganisms in each sediment belonged to unclassified microorganisms. The control sediments (CNTs0%-Crude oil0%) were mostly inhabited by Flavobacteriales (23.7%), Cytophagales (4.4%), Clostridiales (4.4%), Actinomycetales (4.4%), Bacterioidales (4%), Lactobacillales (2.4%), Sphingobacteriales (2.4%), Rhizobiales (2%), Acholeplasmatales (2%) (Fig. 3.4. 2) and many other orders with lower than 2% population (Fig 3.4.6).

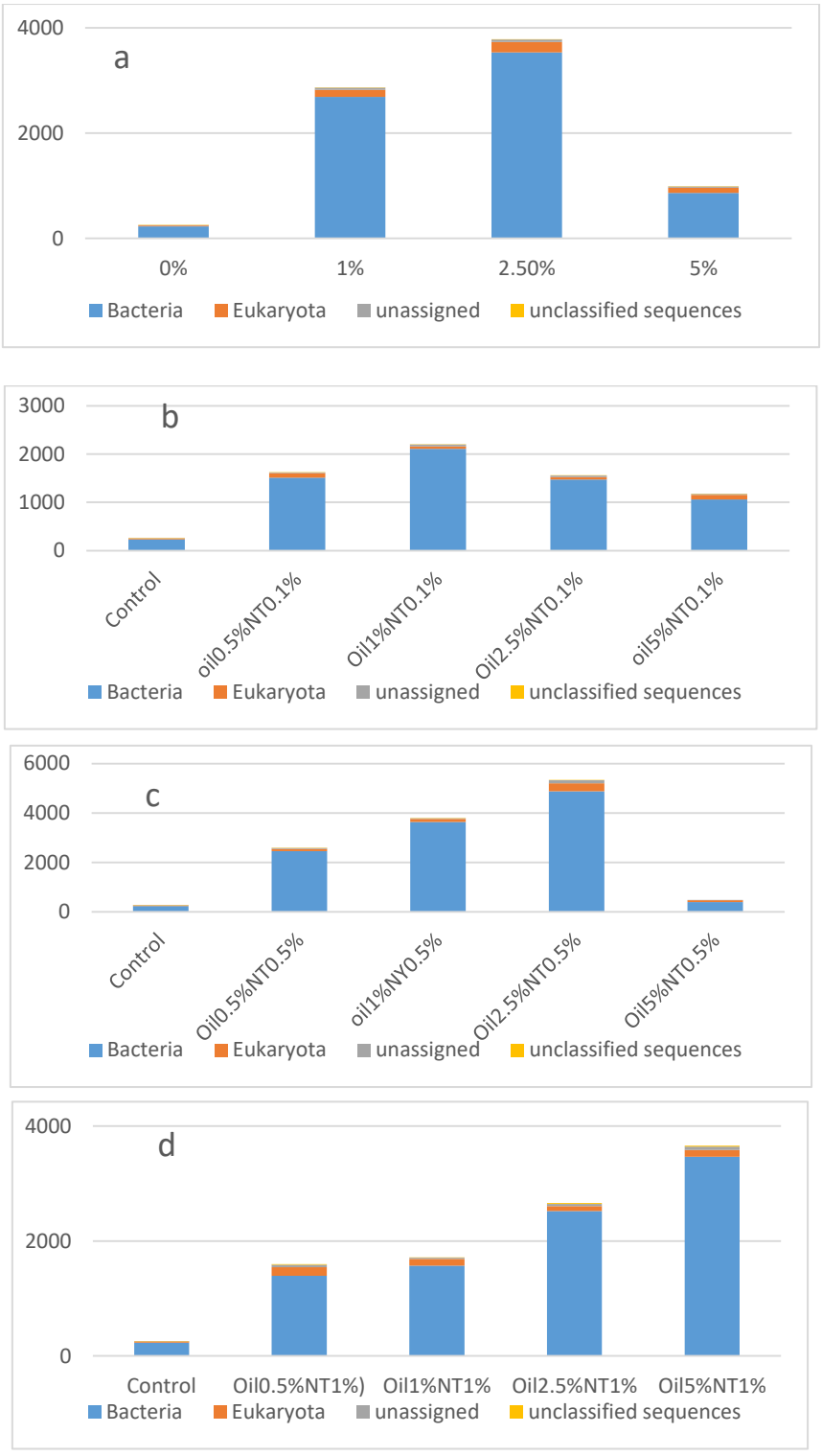


Figure 3.4.1. Total microbial abundances in the sediments treated with different concentrations of carbon nanotubes and crude oil; group a: Crude oil; group b: 0.1% carbon nanotubes, group c: 0.5% carbon nanotubes and group d: 1% carbon nanotubes.

There were increases in the ratios of several orders in the sediments group a, especially Actinomycetales, Bacillales, Bacterioidales, Chroococcales, Melosirales and Nitrospirales (Fig. 3.4.3). Furthermore, many of microorganisms, especially Acholeplasmatales and Flavobacteriales, were susceptible to crude oil treatments. No big differences were seen in the ratio of Clostridiales after treatment with different concentrations of crude oil. In the sediments group a with higher numbers of reads (CNTs0%crude oil 2.5%) Flavobacteriales (21.9%) were still the dominant microorganisms, followed by Clostridiales (15.2%), Bacterioidales (10.6%), Actinomycetales (8.5%), Erysipelotrichales (7.4%), Lactobacillales (4.1%), Cytophagales (3.4%) (Fig. 3.4.3) and many other orders less than 3% ratios (Figure 3.4.6).

These microbial distributions were changed in the sediments group b (CNTs 0.1%), where Flavobacteriales were the dominant microorganisms. In this sediments, the ratios of Actinomycetales, Bacterioidales, Clostridiales, Cytophagales, Lactobacillales, Naviculales, Verrucomicrobiales and many others increased in comparison to the control sediment (Fig. 3.4.3). However, their ratios were certainly dependent on the concentrations of crude oil and CNTs. In the other word, although abundances of some of the orders increased at higher concentrations of the treatments, others could not resist the toxicity of higher doses and their abundances dropped after the primary increases seen at lower concentrations. Since the highest abundances of microorganisms in this group were achieved in the sediment spiked with 1% crude oil and 0.1% CNTs, we focused mainly on this sediment for more analysis. In this sediment, the dominant organisms can be listed as Flavobacteriales (19%), Clostridiales (12.7%), Erysipelotrichales (9.8%), Nostocales (4.2%), Bacterioidales (4.0%), Cytophagales (3.0%), Desulfobacterales (2.1%), Lactobacillales (1.9%), Actinomycetales (1.6%), Verrucomicrobiales (1.4%) (Fig. 3.4.3) and many other orders with lower than 3% ratios (Fig. 3.4.6).

In the group c, there were decreases mainly in the ratios of Flavobacteriales, Actinomycetales, Chroococcales, Melosirales, Rhodobacterales, Sphingobacteriales and many other organisms (Fig. 3.4.4). At the same time, there were increases in the ratios of Bacterioidales, Clostridiales, Erysipelotrichales, Cytophagales, Lactobacillales, Naviculales, Nostocales, Verrucomicrobiales and many other organisms. The dominant organisms in the sediments treated with 2.5% crude oil and 0.5% CNTs can be listed as Flavobacteriales (15.6%),

Erysipelotrichales (8.9%), Clostridiales (7.5%), Bacterioidales (5.3%), Actinomycetales (3.9%), Cytophagales (3.4%), Naviculales (3.3%) (Fig. 3.4.4) and many others with less than 3% abundance (Fig. 3.4.6).

Ultimately, in the group d (treated with 1% CNTs) there were increases in the ratios of Bacterioidales, Clostridiales, Cytophagales, Erysipelotrichales, Lactobacillales, Naviculales, Verrucomicrobiales and many other organisms (Fig. 3.4.5). However, the ratios of Flavobacteriales, Chroococcales, Melosirales, Rhodobacterales, Sphingobacteriales and many other orders decreased in these sediments in comparison to the control. Microbial profile in the sediments treated with 1% CNTs and 5% crude oil changed to Clostridiales (13.2%), Flavobacteriales (11.5%), Erysipelotrichales (7.7%), Bacterioidales (7.3%), Actinomycetales (4.6%), Lactobacillales (3.3%), Cytophagales (2.6%) (Fig. 3.4.5) and many other orders less than 3% ratios (Fig 3.4.6).

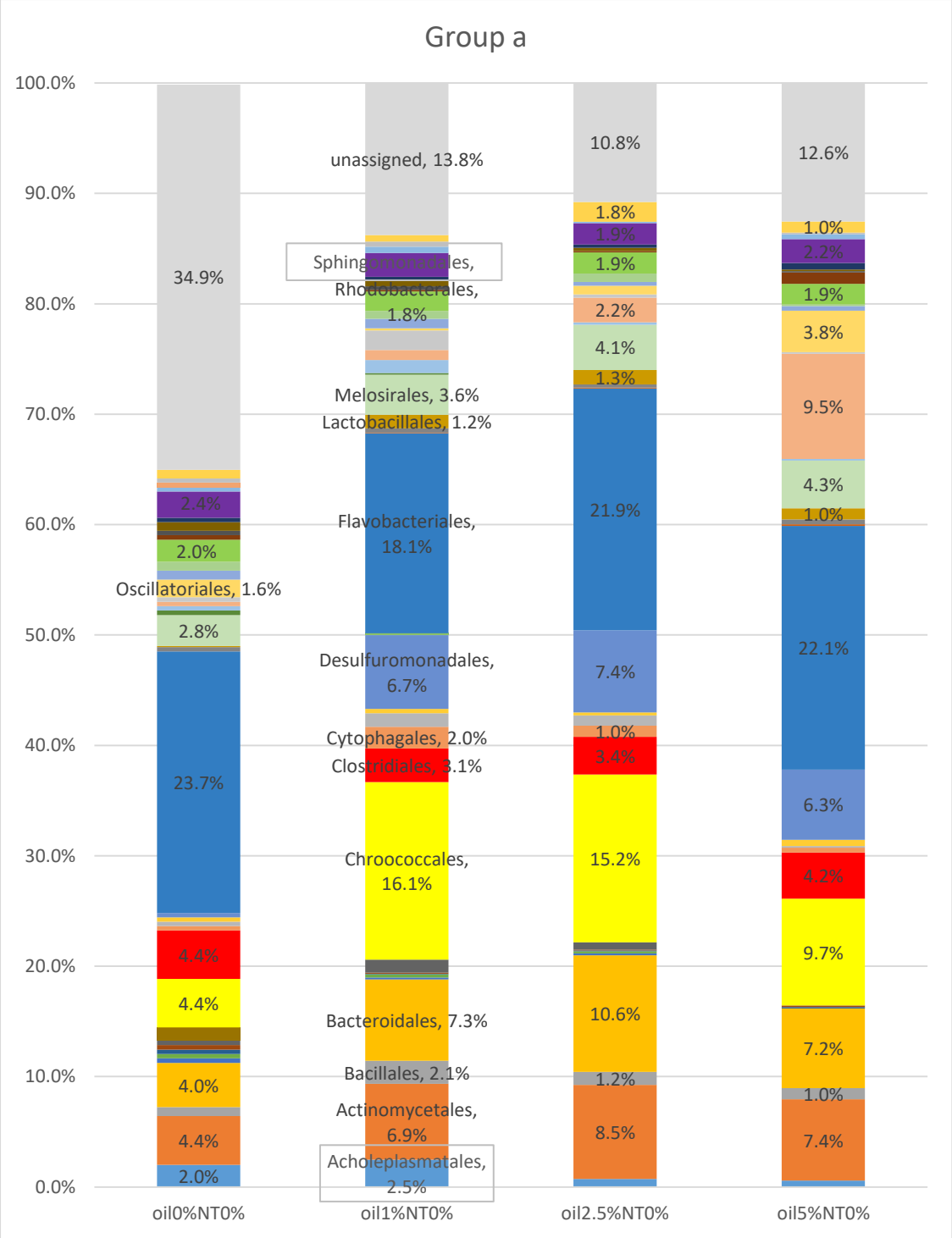


Figure 3.4.2 The ratios of microorganisms in the control sample and the samples treated with different concentrations (0.5%, 1%, 2.5% and 5%) of crude oil (group a). Each order in these columns are characterised with a specific colour.

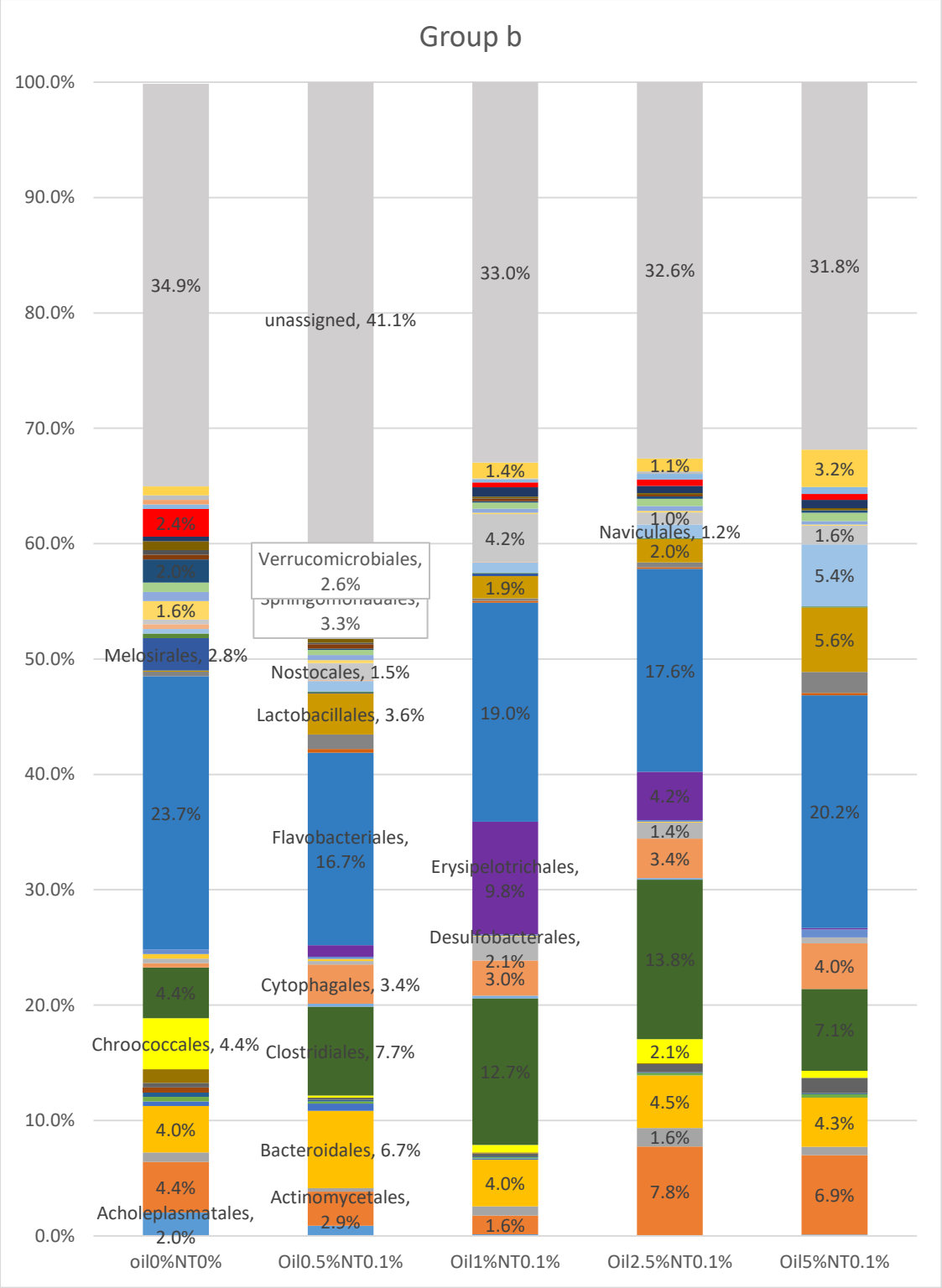


Figure 3.4.3. The ratios of microorganisms in the control sample and the samples treated with 0.1% carbon nanotubes (group b) and different concentrations (0.5%, 1%, 2.5% and 5%) of crude oil. Each order in these columns are characterised with a specific colour.

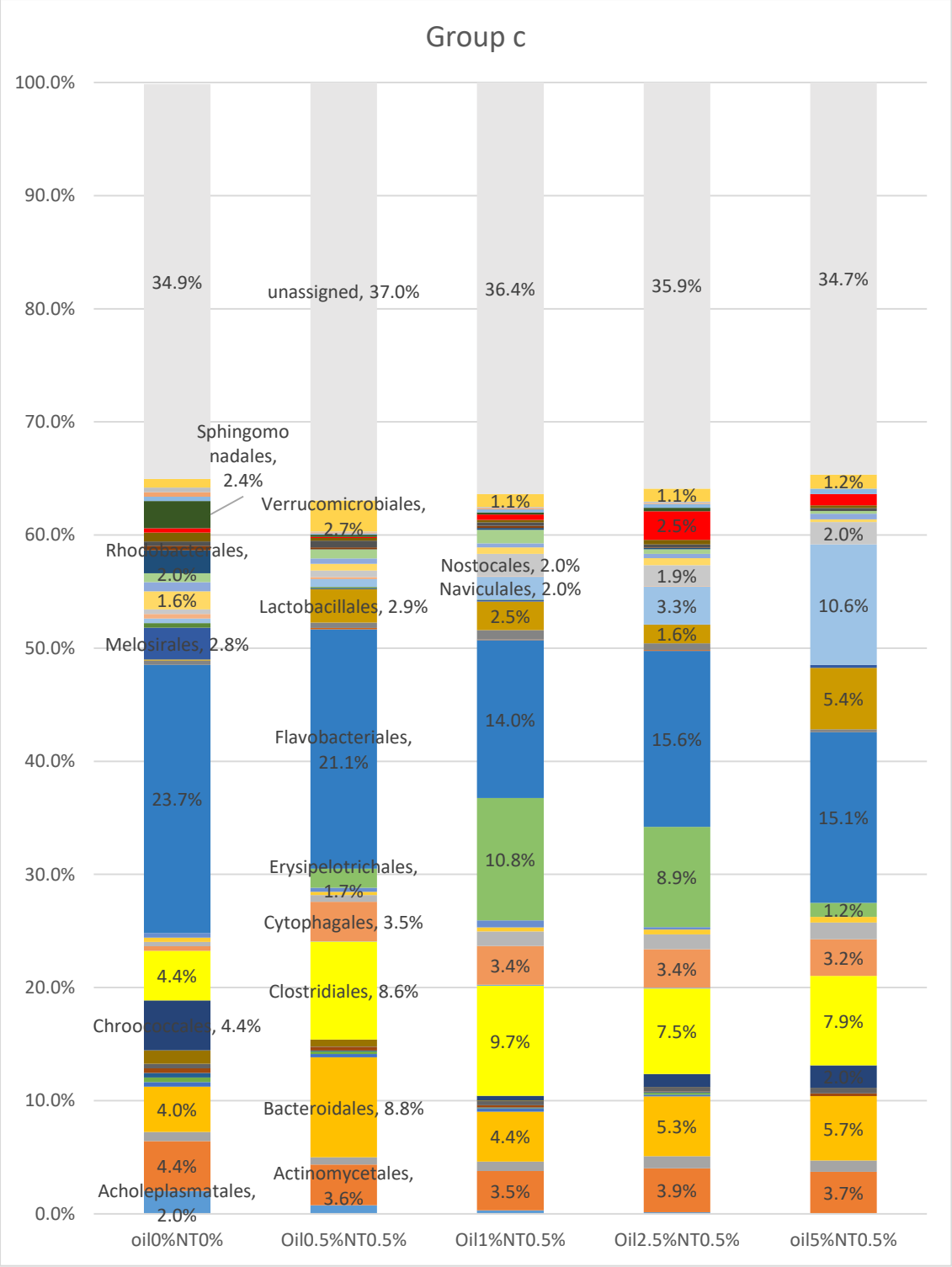


Figure 3.3.4 The ratios of microorganisms in the control sample and the samples treated with 0.5% carbon nanotubes (group c) and different concentrations (0.5%, 1%, 2.5% and 5%) of crude oil. Each order in these columns are characterised with a specific colour.

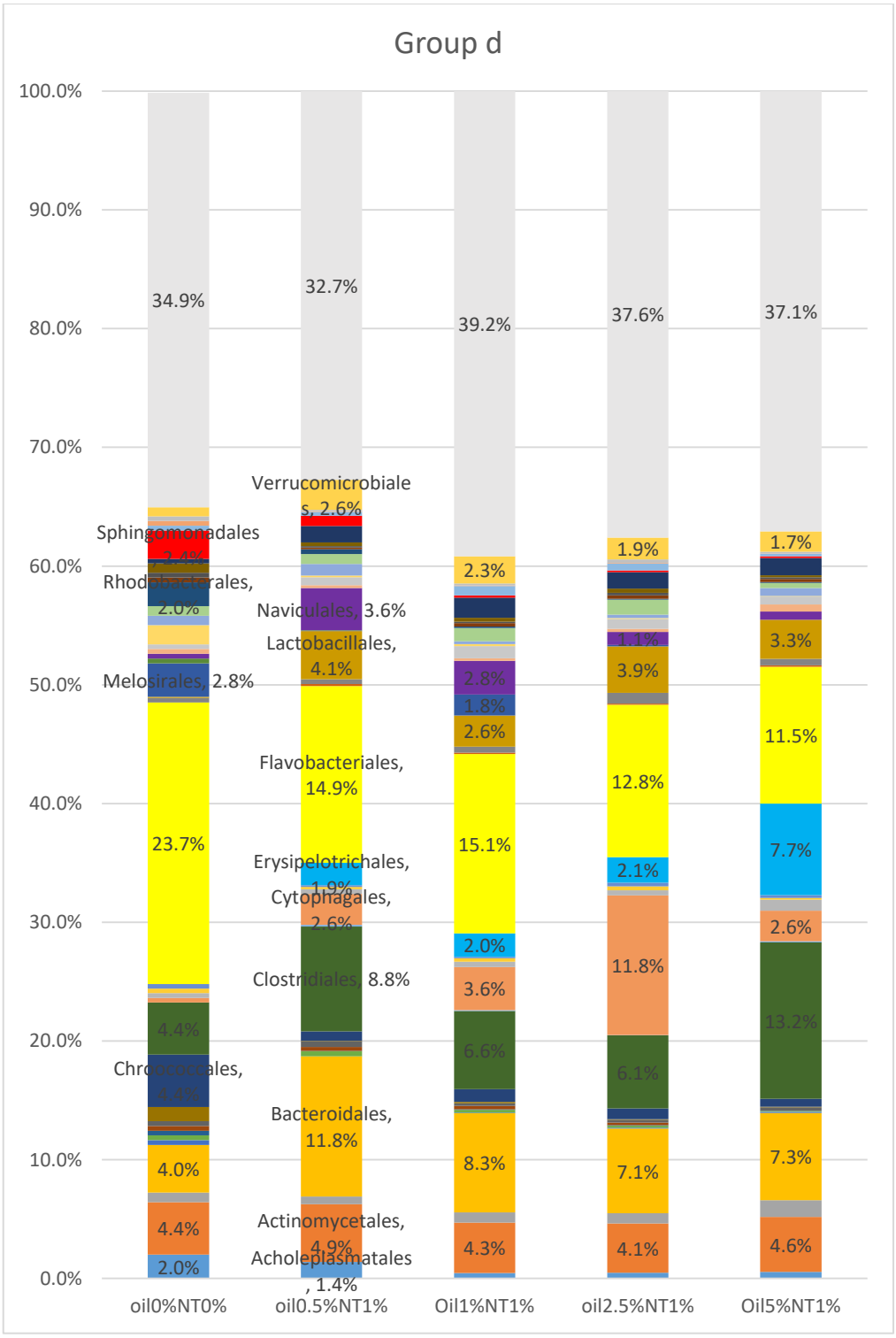


Figure 3.4.5 The ratios of microorganisms in the control sample and the samples treated with 1% carbon nanotubes (group d) and different concentrations (0.5%, 1%, 2.5% and 5%) of crude oil. Each order in these columns are characterised with a specific colour.

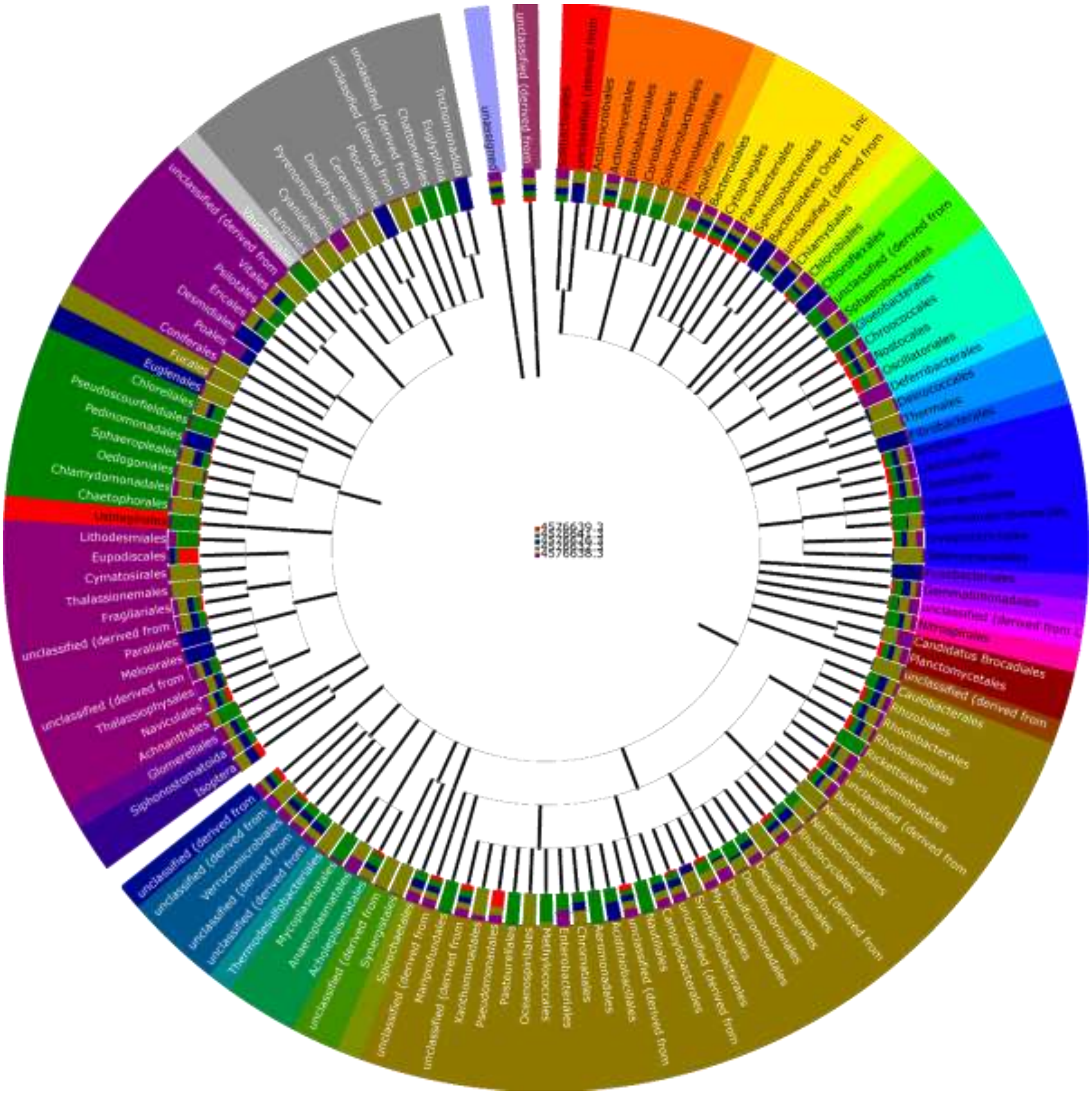


Figure 3.4.6. List of the microbial orders in the sediments control and the sediments from group a, b, c and d with higher numbers of reads (F38, F41, F25, F31 and F37).

3.4.5 Discussion

Microbial biomass is a critical factor in the bioremediation of environmental pollutants, including crude oil. By now, several strategies, such as bioaugmentation and bio-stimulation have been used to increase the microbial biomass in the sites contaminated with crude oil (Sayara et al., 2011, Sprocati et al., 2012). However, preparation of the microbial biomasses at industrial scales in order to be used for bioaugmentation is very expensive (Megharaj et al., 2011). Furthermore, since the mixture of microorganisms used in industry for remediation of

crude oil is not adapted to different physicochemical conditions, such as extreme temperatures (Conant et al., 2011, Rao, 2011), high salinity (McGenity, 2010), low or high pH (Riser-Roberts, 1998), they are not applicable in all conditions. In addition, the supplementary nutrients used for microbial growth, including nitrogen and phosphorous containing fertilizers, are not cost effective when they are used at industrial scales (Abbasian et al., 2015a). As a result, employment of the indigenous microorganisms for degradation of hydrocarbons in the crude oil-contaminated sites seems an appropriate strategy to remove these contaminants. Although many microorganisms are able to degrade hydrocarbons, the toxicity effects of several hydrocarbons present in crude oil on the microbial population (Tang et al., 2011, Hubálek et al., 2007) and the insolubility of these compounds in water (Morgan and Watkinson, 1994, Sánchez-Palencia González, 2011) are two restrictive factors for microbial growth in these conditions. In our previous study (unpublished work), we showed the ability of multiwall carbon nanotubes to increase the microbial abundances in the environments. Hence, we hypothesised that the employment of carbon nanotubes can increase the total abundances of microorganisms in the crude oil contaminated sites. Since these microorganisms are present in the crude oil contaminated site, they may be useful for bioremediation of these compounds.

Based on results of this study, the mixtures of multiwall carbon nanotubes and crude oil affected the total abundances of microorganisms in the sediments. However, the effects depended on the concentrations of both treatments. The peak of total microbial abundance in the sediments group a, where 0.1% MWCNT was added to the sediments, was achieved at low concentration (1%) of crude oil treatment. However, when higher doses (0.5% and 1%) of the multiwall carbon nanotubes were used in these sediments, the total abundances of microorganisms were increased at higher concentrations of crude oil. Therefore, the more concentrations of MWCNTs was used, the more microbial abundance was achieved at higher concentrations of crude oil contamination. As a result, amendment of multiwall carbon nanotubes can increase the microbial abundances in crude oil-contaminated sites, which is one of critical aspects of bioremediation.

In addition to the total abundances of microorganisms, the microbial diversity and the ratio of each microbial population were affected by the combinations of crude oil and multiwall carbon nanotubes. Several microorganisms, such as Flavobacteriales, Acholeplasmatales,

Burkholderiales, Chlamydomonadales, Chlorellales, Chromatiales, Desulfovibrionales, Gemmatimonadales and Myxococcales were highly susceptible to treatment with both crude oil alone and mixtures of crude oil and carbon nanotubes. Although the ratio of Flavobacteriales in the sediments treated with crude oil and multiwall carbon nanotube decreased in comparison to the control sediment, this order, especially *Flavobacterium columnare*, was still the dominant microorganism in the sediments spiked with different concentrations of crude oil and carbon nanotubes. This order along with Sphingomonadales and Bacterioidales belong to a group of Gram negative anaerobic chemoheterotrophic bacteria referred to as phylum Bacteroidetes (Whitman and Parte, 2009). Although there were no significant changes in the ratio of Sphingomonadales, the ratio of Bacterioidales, especially *Porphyromonas gingivalis*, were increased in the sediments treated with the mixtures of crude oil and carbon nanotubes. In both cases, increases in the ratios of these microorganisms or the absence of a decrease in this value is indicative of the ability of these bacteria to survive in crude oil contaminated environment and probably to degrade the hydrocarbon molecules in crude oil. The presence and activity of the members of these orders in sediments contaminated with hydrocarbons has been proven in previous studies (Prince et al., 2010, Gerdes et al., 2005, Evans et al., 2004). Although *Porphyromonas gingivalis* is usually known as an oral pathogen (Medina et al., 2014), its presence in crude oil sources has been reported as well (Shimoyama et al., 2009).

The degree of susceptibility of microorganisms in these sediments depended mainly on the concentrations of both crude oil and carbon nanotube treatments. Actinomycetales, for instance, are an example where treatment with high concentrations of crude oil alone and with mixture of crude oil and low concentrations of carbon nanotubes could increase their ratios in comparison to the control sample. The ratios of this order at higher concentrations of carbon nanotubes and crude oil did not reach levels lower than the control level. Clostridiales, Erysipelotrichales and Lactobacillales were other Gram positive bacteria where their ratios increased in the sediments treated with both MWCNTs and crude oil, and these values increased at higher concentrations of treatments. The highly abundant presence of Clostridiales (Gieg et al., 2008, Sherry et al., 2013, Mbadinga et al., 2011), Actinomycetales (Akbari and Ghoshal, 2015), Erysipelotrichales (Wang et al., 2014b) and Lactobacillales (Al-Mailem et al., 2015) has been reported in several studies working on the biodiversity of microorganisms in crude oil contaminated sites and/or their activity in biodegradation of crude oil or bioenergy production through degradation of residual oil.

Naviculales, mainly *Anabaena* sp., and Nostocales, mainly Nostoc sp. were two phototrophic microorganisms with the ability to survive in the sediments treated with only low concentrations of crude oil. However, treatment with mixtures of crude oil and carbon nanotubes could increase their ratios at higher concentrations of crude oil contamination. The increases in the abundances of *Anabaena* sp., belonging to order Naviculales, and Nostoc sp., belonging to Nostocales, in the presence of crude oil contamination was shown in several studies (Raghukumar et al., 2001, Zhao et al., 2011, Hirano et al., 2004). Furthermore, Gamila et al. (Cai et al., 2011) showed that the biomass of *Anabaena* sp. increased with exposure to crude oil and the majority of the C10–C24 n-alkanes (99.5%) and PHAs (97.5%) were removed in the sites contaminated with crude oil. Although these algae are not involved in the direct degradation of hydrocarbons, they encourage the growth and activity of crude oil-degrading chemoheterotrophic bacteria through the production of oxygen, nitrogen fixation and production of simple organics (Papizadeh et al., 2011). Therefore, increases in the ratios of phototrophic organisms using carbon nanotubes can be a promising technique to improve the bioremediation of the sites contaminated with crude oil.

Overall, most of the dominant microorganisms in the sediments treated with combinations of crude oil and MWCNTs were also present in the sites contaminated with crude oil alone. However, the ratios of these microorganisms in the sites contaminated with mixture of crude oil and MWCNTs were higher than the sites contaminated with crude oil alone. Therefore, the treatments with mixture of crude oil and carbon nanotubes could increase both the numbers of effective orders (as was proven based on Sharon's index) and their ratios in the sediments in comparison to the untreated control and the sediments treated with only crude oil. These positive effects of carbon nanotubes on microbial populations could be due to the aggregation of nanotubes and therefore serving as microenvironments for bacterial attachment and growth (Koelmans et al., 2009, Kang et al., 2008). Furthermore, attached bacteria are known to be more protected from predation compared to free cells (Fuhrman et al., 2015).

As is discussed here, based on the literature the majority of the dominant microorganisms in these sediments treated with mixture of crude oil and carbon nanotubes are able to utilise the hydrocarbon molecules present in the crude oil. This means not only that the treatment with

multiwall carbon nanotube did not destroy the microorganisms useful in oil bioremediation, but could increase their growth in contaminated sites. As a result, these particles can be applicable for enhancing microbial growth, especially in the in situ experiences where employment of indigenous microorganisms is critical for a special activity such as removal of the contaminants. However, given our focus in this study was confined to the evaluation of the influence of carbon nanotubes on the microbial abundance and their diversity over a shorter period of time, further studies are required to see the effect of nanotube induced enhanced microbial growth on the remediation of crude oil in terms of removal of oil content.

3.5 Microbial diversity and hydrocarbon degrading gene capacity of a crude oil field soil as determined by metagenomics analysis

Published in: Biotechnology Progress, DOI: 10.1002/btpr.2249

3.5.1 Abstract

Soils contaminated with crude oil are rich sources of enzymes suitable for both degradation of hydrocarbons through bioremediation processes and improvement of crude oil during its refining steps. Due to the long term selection, crude oil fields are unique environments for the identification of microorganisms with the ability to produce these enzymes. In this metagenomic study, based on Hiseq Illumina sequencing of samples obtained from a crude oil field and analysis of data on MG-RAST, Actinomycetales (9.8%) were found to be the dominant microorganisms, followed by Rhizobiales (3.3%). Furthermore, several functional genes were found in this study, mostly belong to Actinobacteria (12.35%), which have a role in the metabolism of aliphatic and aromatic hydrocarbons (2.51%), desulfurization (0.03%), element shortage (5.6%) and resistance to heavy metals (1.1%). This information will be useful for assisting in the application of microorganisms in the removal of hydrocarbon contamination and/or for improving the quality of crude oil.

3.5.2 Introduction

Soil contamination with crude oil is a serious concern for human and ecological health (Jin et al., 2012). Annually millions of tons of crude oil are extracted to supply energy and primary materials for industries, and this demand has been increasing as a result of growth in population and industries (Abbasian et al., 2015a). In addition to the natural release of hydrocarbons, several industrial activities, including crude oil extraction, transportation, refining and utilisation, especially in automobile service stations, cause environmental pollution (Abbasian et al., 2015a). This situation deteriorates with time when the components are exposed to weathering processes, which lead to increases in the relative levels of higher molecular weight hydrocarbons with higher density and viscosity (Pietroski et al., 2015). Although several mechanical and chemical techniques are available for cleaning these contaminants from the environments, in situ bioremediation is a very efficient and cost effective approach (Yoon, 2015). Since the use of native microorganisms for *in situ* removal of these contaminants has

shown promise in several studies (Yoon, 2015), an investigation of the diversity of microorganisms in contaminated areas and of the effects of environmental conditions on the degradation ability of these organisms can improve the management of the bioremediation rate in these sites. Although many studies have been performed to investigate the microbial diversity and the genes involved in degradation of aliphatic and aromatic hydrocarbons (Suja et al., 2014, Uhlik et al., 2013), most of these studies focussed on a special community of microorganisms or on a limited number of genes of interest. Furthermore, these studies mostly focussed on the genes involved in the degradation of hydrocarbons or the production of surfactants, and less consideration has been given to other genes effective in bioremediation, such as heavy metal resistant genes and the genes required for acquisition of essential nutrients. In addition, since the levels of nitrogen, phosphorus, sulfur and iron of the crude oil contaminated sites are very low due to the predominance of hydrocarbons (Uhlik et al., 2013), investigations of the genes involved in the acquisition of these nutrients can also be useful in the development of optimized physical, chemical and environmental conditions that may improve rate and efficiency of biodegradation process.

Since microbial culture is limited to the isolation and identification of limited numbers of microorganisms, this technique is not suitable for investigation of whole microbial diversity and their genomes of all microorganisms present in a particular environment (Abbasian et al., 2015b). However, recent metagenomic approaches enable researchers to analyse the whole microbial diversity, and the genetic capacity for the active metabolic pathways present in a given environment (Gong et al., 2013). This technique is especially efficient for the study of the biodiversity and genome analysis of complex environmental samples where most of microorganisms cannot be cultured under normal laboratory conditions (Abbasian et al., 2015b). Moreover, this system enable operators to establish a correlation between microbial diversity and the level of hydrocarbon(s) in a contaminated site (Abbasian et al., 2015b). Up to now, several studies have been performed on the microbial communities and metagenomic capacities of soils contaminated with crude oil and its derivatives (Liang et al., 2011, An et al., 2013, Yergeau et al., 2012b). These studies have revealed large variations in the biodiversity and abundances of microorganisms in different geographical locations. Furthermore, big differences were observed in the genetic capacities of microbial communities identified at each location. Therefore, studies such as these give excellent opportunities for finding new microbial strains and the genes involved in both bioremediation of hydrocarbon contaminants and improving crude oil quality through refining biological process. In this study, the whole

microbial diversity of a crude oil contaminated field soil and the genes involved in the degradation of hydrocarbons and microbial adaptations to this environment was investigated using a metagenomic approach.

3.5.3 Material and methods

Overall, 10 kg surface (0-10 cm depth) soil were sampled from different areas at the vicinity of crude oil wells (Perth, Western Australia, Australia). Contaminated soil samples were transported from the field to the laboratory under standard refrigerated (4^{o} C) conditions. To ensure the quality of the work, following sifting the soil using a 2 mm metal sieve, the soil samples were thoroughly mixed and the sampling for DNA extraction was performed from different parts of the soil samples. The soil characteristics were measured according to the protocols described in text books (Sparks, 2003, Rayment and Higginson, 1992) as 6.8 pH at room temperature, 2.8% moisture and 612 µS conductivity.

3.5.3.1 TPH analysis

Following extraction of TPH compounds from soil samples using the sequential ultrasonic solvent extraction method (Risdon et al., 2008), the level of contamination was quantified by a gas chromatograph fitted with a flame ionization detector (GC-FID, Agilent 6890). The Chromatography process was conducted on a fused-silica capillary column BPX-5 from SGE (15 m × 0.32 mm internal diameter) coated with HP-5 (0.10-µm film thickness). Helium (2.5 mL min^{-1}) was used as the carrier gas, and the FID detector was adjusted at 300°C. Splitless injection with a sample volume of 1 µL was applied and the oven temperature was increased from 50 °C to 300 °C at a gradient of 25 °C min^{-1} and held at this temperature for 5 min. Overall, the total run time was 19.6 minutes. The quantification of hydrocarbons were performed by Agilent Chemstation Software through integration and calibration of peaks of a standard concentration of an external calibration standard namely Hydrocarbon Window Defining Standard (C_8-C_{40}) from AccuStandard (Miller and Miller, 1987). Overall, five concentrations of external calibration standard in the range expected in the samples were analysed, and a linear curve fit with a R^2 value of 0.997 was obtained. The CCV (Continuing Calibration Verification) was analysed at the start and end of every 20 samples and CCV recovery was 95-110% of true value. Hexane was run as blank with every 10 samples to test for cross contamination. The surrogate (*o*-terphenyl) was spiked at a level to produce a recommended extract concentration of 20 µg/mL. Surrogate recoveries ranged 70-120% for all

the samples, and 25 mg kg^{-1} was the minimum concentration of TPH detected (MDL) through this analytical method. To confirm the quality of the TPH analysis, the soil was vortexed properly and four samples were taken for this analysis.

3.5.3.2 DNA sequencing using Illumina Hiseq platform and data analysis

To confirm the quality of DNA sequencing, the soil was vortexed properly and two samples were taken for metagenomic analysis. The DNA from the soils was extracted using a power soil DNA kit (MO BIO), according to manufacturer's instructions. First, after a gentle vortex of 0.25 g sediment samples, the contents were mixed with 60 µl solution C1. Following a vigorous vortex step for 10 minutes and a centrifuge at 10,000 x g for 30 S, the supernatants (400-500 µl) were mixed up with 250 µl solution C2 in clean 2-ml tubes. These tubes were shaken gently for 5 S and were passed a 4 °C incubation for 5 minutes. After separation of the supernatant by centrifugation at 10,000 x g for 60 S, up to 600 µl supernatants were poured into a new 2-ml tubes, where were mixed with 200 µl solution C3. Once again, there was an incubation period at 4 °C for 5 minutes and a centrifugation at 10,000 x g for 60 S. A total of up to 750 µl of the resulting supernatants were mixed up with 1200 µl of solution C4 in new clean 2-ml tubes. Following a gentle vortex, a portion of these mixtures (675 µl each time) were loaded onto a spin filter, where centrifugation force at 10,000 x g for 60 S made the mixture pass through the filters and DNA stuck on. This step continued to centrifuge all the mixture. These filters were wash out by adding 500 µl solution C5 and a centrifugation at 10,000 × g for 30 seconds. After transfer of the spin filters into a new 2-ml tube and addition of 100 µl solution C6, a centrifugation step at 10,000 × g for 30 seconds eluted the DNAs into the solution buffer. The levels of DNAs were quantified by a quantifluor dsDNA system (Promega) and were adjusted to a minimum of 100 ng/µl. The DNA samples were then sequenced on the Illumina Hiseq platform by the Australian Genome Research Facility (AGRF).

The raw data obtained from Illumina platform were submitted directly in Metagenome Rapid Annotation using Subsystem Technology (MG-RAST) (Glass and Meyer, 2011), where all the data was trimmed by removing uninformative and/or duplicative reads before any further studies. The microbial diversity of the data was determined using M5RNA annotation source

(minimum percentage identity cut off of 97%, an E-value cutoff of 1e-5 and minimum alignment length cutoff of 50 bp). Furthermore, the M5NR annotation source based on a minimum percentage identity cut off of 70%, an E-value cutoff of 1e-5 and a minimum alignment length cutoff of 15 bp was used to analyse the best hit classification of genes in this data set. In addition, the hierarchical classification of the genes and their functional abundances were determined through KO (KEGG orthology), NOG (non-supervised orthologous groups) and COG (cluster of orthologous groups) subsystem sources and based on minimum percentage identity cut off of 70%, E-value cutoff of 1e-5 and cutoff of 15 bp. Further statistical analysis were performed using the excel program.

3.5.3.3 Nucleotide sequence accession numbers

This metagenomic analysis of crude oil field soil was registered in NCBI under ID SUB713359, BioSample accession SAMN03106209 and BioProject ID PRJNA263754, and the reads are publicly available in the MG-RAST system as A11 and A12 under project identifiers 4 4582710.3 and 4582711.3.

3.5.4 Results

3.5.4.1 Gas chromatograph assays

The gas chromatograph assays showed that this soil is highly polluted with many different types of hydrocarbons, including >C16-C34 (33317 ppm), >C10-C16 (18520 ppm) and >C34-C40 (30 ppm), and its TPH (Total Petroleum Hydrocarbon) level is largely high (51867 ppm).

3.5.4.2 Statistics of reads produced by HiSeq Illumina

The dataset obtained from the sequencing of DNA from the crude oil well field (A21 and A22) were uploaded to the online metagenomic analysis tool, MG-RAST (*metagenomics* Rapid Annotation using Subsystem Technology) and were automatically filtered by QC software. These two files contained 30656512 fragments with an average length of 138 ± 25 bp and 128 ± 32 bp (respectively for A21 and A22) and totally 4920764700 base pairs, among them 1259670862 reads (25%) did not pass the QC pipeline. Overall, 23221149 (75%) and 337603 (1.1%) fragments contain predicted proteins and ribosomal RNA genes, respectively, of which

7872294 (25.7%) and 4743 (0.015%) reads belonged to identified protein and rRNA features. The total GC% value in these dataset was 63 ± 9 %.

3.5.4.3 The microbial communities in the crude oil well field

The microbial orders identified in these samples are listed in Figure 3.5.1. Regardless of the 7.4% of organisms that were unclassified, the majority of the OTUs (84.2%) identified from the data belonged to bacterial phyla, followed by eukaryota (7.5%), archaea (0.8%) and viruses (0.04%) (Fig. 3.5.2). Among the bacterial phyla, the crude oil field soils community was comprised mostly of Proteobacteria (32.1%), Actinobacteria (12.3%), Firmicutes (5.8%), Acidobacteria (2.5%), Bacteroidetes (2.3%), Planctomycetes (2.1%), Gemmatimonadetes (1.6%) and many other bacterial phyla with less than 1% abundance (Fig. 3.5.2). At the order level, Actinomycetales (9.8%), Rhizobiales (3.3%), Sphingomonadales (2.9%), Burkholderiales (2.6), Clostridiales (2.5%) and many other orders lower than 2% abundance were the dominant OTUs (Fig. 3.5.2). Fungi 2.95% (mostly Glomerellales) and many small animals from different orders (4.32%) were classified as other and were not considered for further studies.

3.5.4.4 Classification of microorganisms based on their metabolism and respiration

Following information in the literature (Whitman and Parte, 2009, Schink et al., 2011) and based on microbial ability to supply their carbon and energy sources, the known microbial orders found in this study were categorized into four different metabolic classes (photoautotrophic microorganisms, photoheterotrophic bacteria, chemolithotrophic bacteria and chemoheterotrophic bacteria). Since some particular species of microorganisms in an order can show a completely different metabolic pathway in comparison to other strains from the same order, the properties of each strain were analysed separately. Based on this analysis, the majority (88%) of these microorganisms belonged to chemoheterotrophic bacteria, followed by photoautotrophic microorganisms (3%), photoheterotrophic bacteria (5%) and chemolithotrophic bacteria (4%) (Fig. 3.5.3). Most of the chemoheterotrophic bacteria in this field belonged to aerobic chemoheterotrophic bacteria (62%), followed by facultative aerobic/anaerobic bacteria (23%) while absolute anaerobic bacteria consisted only 15% of this community (Fig. 3.5.4).

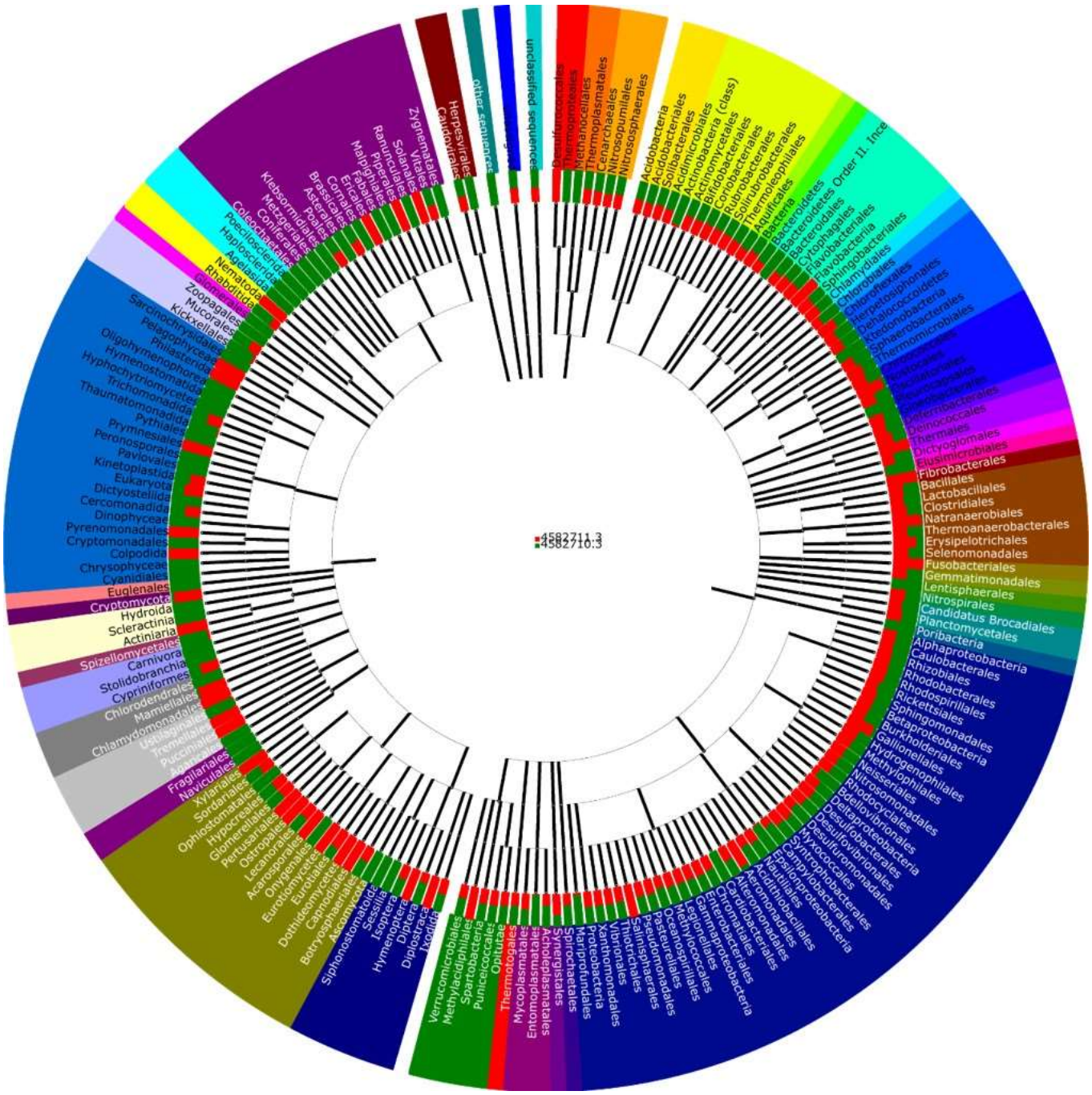


Figure 3.5.1. The microbial diversity in the crude oil field; each colour indicates a special phylum.

Figure 3.5.2. Microbial diversity and their ratios in the crude oil field; the orders with lower than 1.0% are not considered.

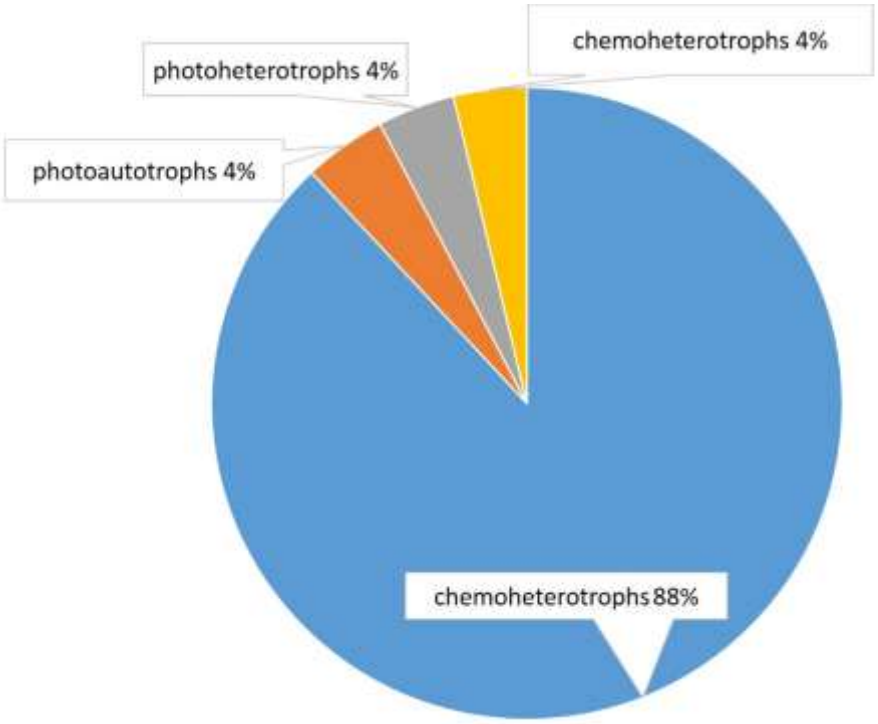


Figure 3.5.3. The ratios of microorganisms identified in the crude oil field based on the Carbon/energy requirem

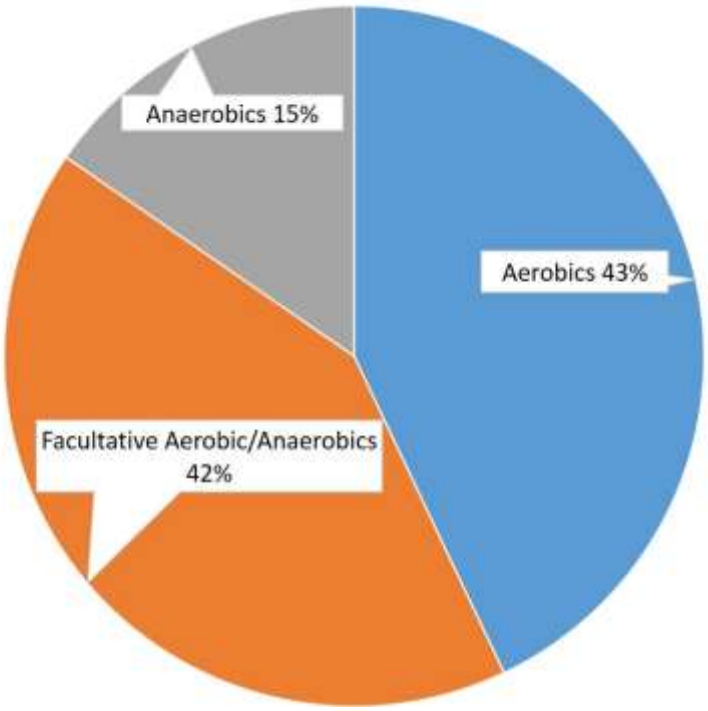


Figure 3.5.4. The ratios of microorganisms identified in the crude oil field based on O_2 requirement.

Based on analysis of the data using the M5NR annotation source, 98.2% of the total genes found in the soil samples collected from crude oil well field belonged to bacteria, followed by Eukaryota (0.7%), Archaea (0.7%) and viruses (0.3%) (Fig. 3.5.5). Most of these functional genes (49.14%) belonged to Proteobacteria, followed by Actinobacteria (12.35%), Planctomycetes (4.07), Acidobacteria (3.4%), Bacteroidetes (3.1%), Firmicutes (2.04%) and many other organisms with less than 2% abundance (Fig. 3.5.5). Overall, these genes could be categorized as being involved in several metabolic pathways, including clustering (15.1%), carbohydrate metabolism (10.8%), amino acid metabolism (9.4%), protein metabolism (8.9%), miscellaneous (7.2%), cofactors and vitamins (6.8%) and many others including housekeeping genes at less than 5% (Fig. 3.5.6). The genes involved in the metabolism of aromatic hydrocarbons (1.8%) were sub-classified into the genes involved in aerobic (14%) and anaerobic (18%) pathways as well as the genes involved into peripheral (46%) and intermediate (22%) metabolisms.

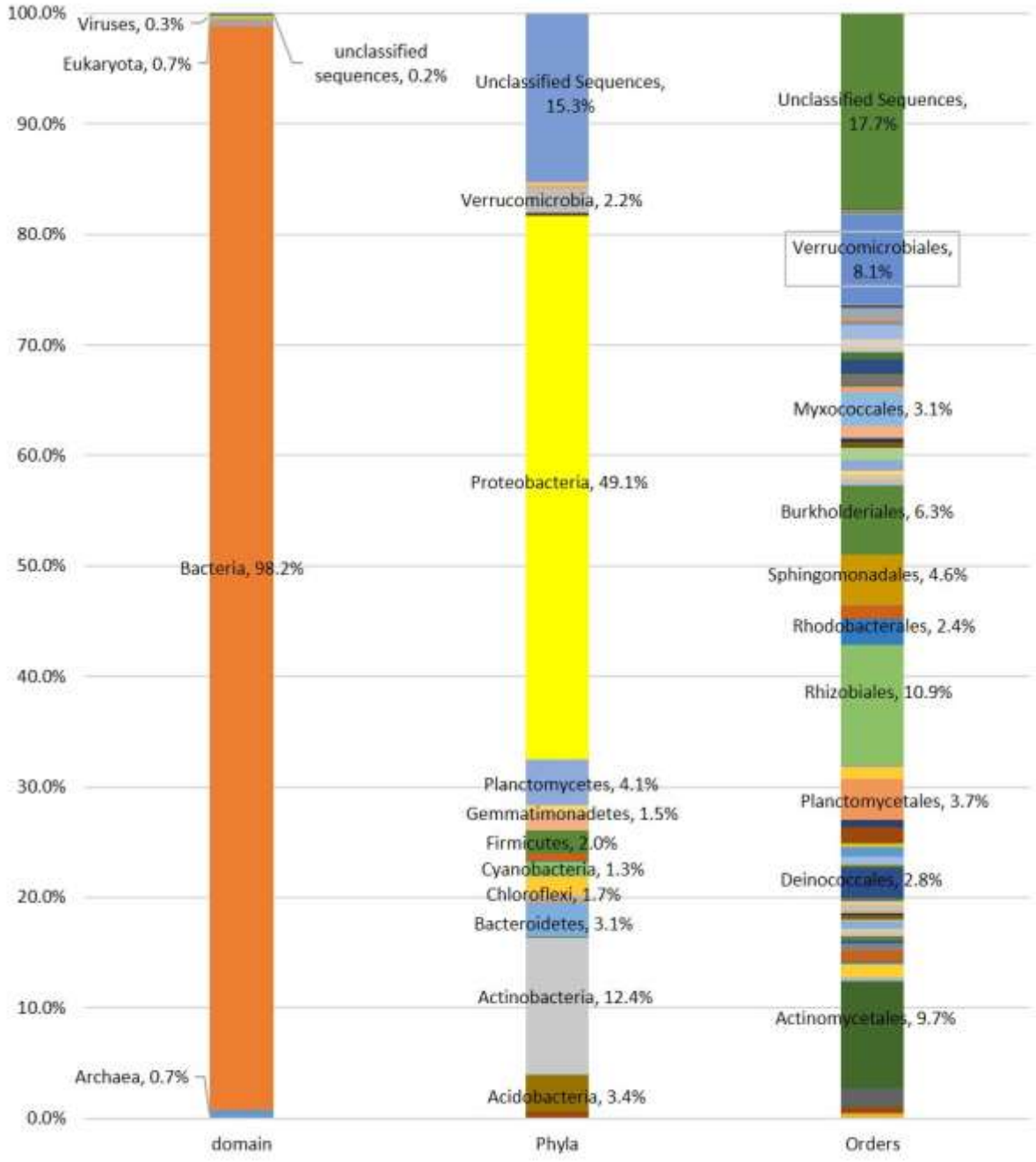


Figure 3.5.5. The ratios of distribution of genes in the microorganisms identified in the crude oil field; the orders with lower than 2.0% are not considered.

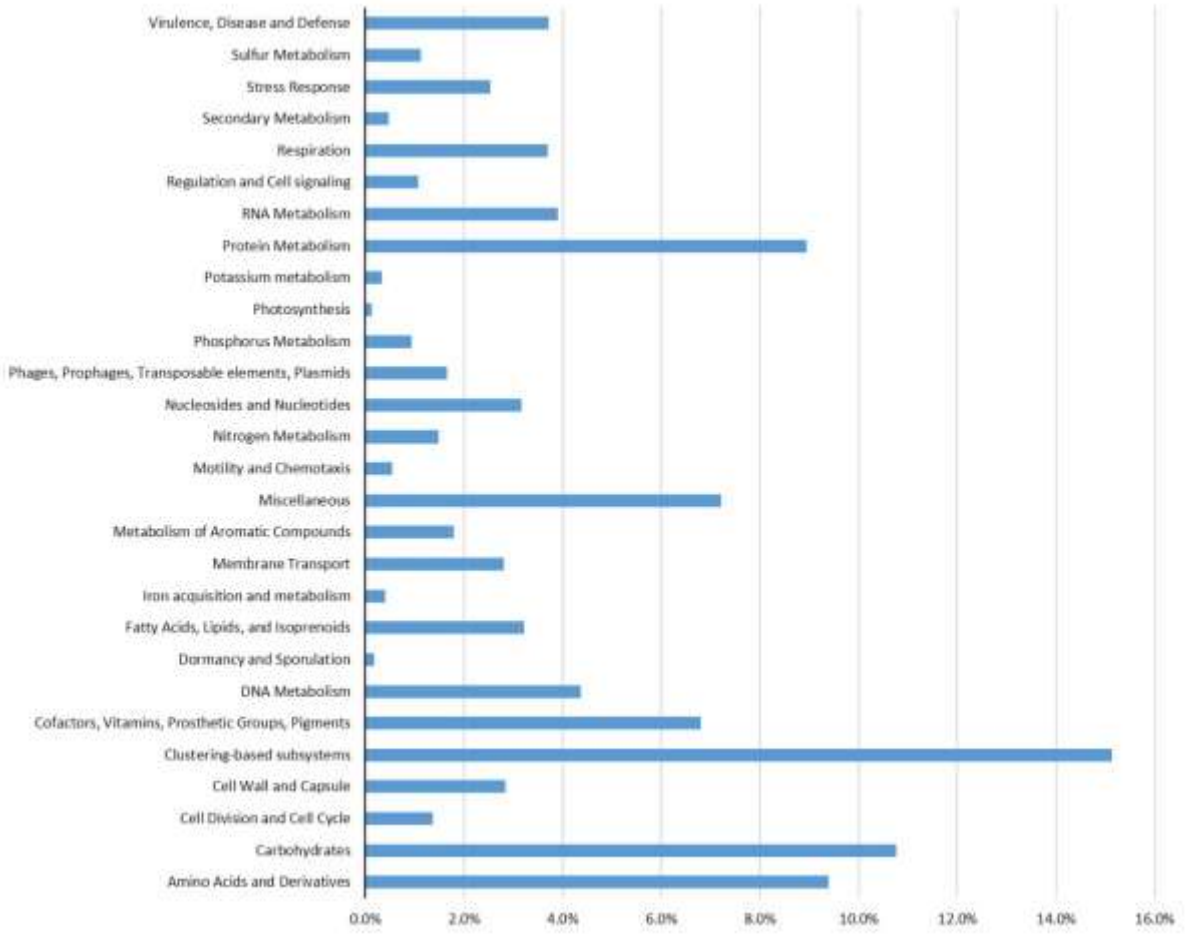


Figure 3.5.6. Distribution of genes in the metagenome study of crude oil field

3.5.5 Discussion

Environmental contamination due to petroleum hydrocarbons is of growing public concern as large amounts of these compounds are being released deliberately or inadvertently into the environment (Abbasian et al., 2015a). Crude oil extraction processes lead to the release of massive amounts of these compounds to soil and water, and therefore, crude oil fields are highly contaminated with several types of aliphatic and aromatic hydrocarbons originating from crude oil (Jin et al., 2012). Although this phenomenon can lead to tragic environmental problems such as contamination of water sources, these areas also become enriched sources of hydrocarbon degrading microbial populations and the corresponding genes for the degradation of hydrocarbon contaminants(Abbasian et al., 2015a). Metagenomics analysis of such environments is a promising approach that allows analysis of these microbial communities and all the functional genes involved in both the adaptation of microorganisms

to crude oil contamination and the degradation of aliphatic and aromatic hydrocarbons (Abbasian et al., 2015b). In this study, a HiSeq Illumina platform was employed to investigate the microbial population of soil samples collected from a crude oil well field located in Perth, Australia.

Based on the results, bacteria (mostly Proteobacteria, Actinobacteria and Firmicutes) were the dominant microorganisms in this field, and eukaryotes and archaea comprised only a small portion of this community. Similar results were reported by Yergeau et al (Yergeau et al., 2012a) and Liang et al (Liang et al., 2011) through metagenomic analysis on the soil samples obtained from Canadian High Arctic Soils and different areas of China, respectively. Several members of order Actinomycetales were identified in this field, among them genera *Mycobacterium* sp., *Geodermatophilus* sp., *Nocardia* sp. and *Nocardioides* sp., were the most dominant microorganisms. While contamination with crude oil could increase the abundance of *Mycobacterium* sp. in the Canadian High Arctic Soils too, the abundances of *Nocardia* sp. and *Nocardioides* sp. were decreased in these samples (Yergeau et al., 2012b), showing the effects of different environmental conditions, such as soil chemistry and climate, on the microbial community.

Overall, 85% of the microbial strains identified in these samples belonged to obligatory and facultative anaerobic bacteria, which was expected because the soil samples were obtained from near the soil surface, exposed to aerobic conditions. For instance, Dongshan et al (An et al., 2013) showed that the majority of microorganisms in the subsurface reservoirs contaminated with crude oil belonged to methanogenic microorganisms, which shows the effects of oxygen tension on the microbial communities. In addition to a few gram positive bacteria, such as *Mycobacterium* sp., *Geodermatophilus* sp., *Nocardia* sp. and *Nocardioides* sp. several gram negative bacteria, such as Rhizobiales (mostly *Rhodopseudomonas* sp., *Bradyrhizobium* sp., *Rhizobium*, *Mesorhizobium* sp. and *Methylobacterium* sp.), Burkholderiales (mostly *Burkholderia* sp. and *Leptothrix* sp.), Rhodobacterales (mostly *Rhodospirillum* sp., *Magnetospirillum* sp., *Rhodobacter* sp. and *Azospirillum* sp.), Myxococcales (mostly *Chondromyces* sp., *Stigmatella* sp. and *Haliangium* sp.), Pseudomonadales (*Pseudomonas* sp.), Xanthomonadales (mostly *Xanthomonas* sp., *Stenotrophomonas* sp. and *Xylella* sp.) and Solirubrobacterales (*Conexibacter woesei* DSM 14684) comprised the majority of strictly aerobic chemoheterotrophic bacteria in the crude oil field. Among these bacteria, Pseudomonadales (Pacwa-Płociniczak et al., 2014),

Burkholderiales (Dobslaw and Engesser, 2014), Xanthomonas sp. (Salam et al., 2014) and Rhodobacterales (Auffret et al., 2014) are the most frequent microorganisms used for bioremediation and removal of hydrocarbon contamination. However, obligate anaerobic bacteria comprised a small portion of this microbial community and among them, Clostridiales (2.5%) (mostly *Clostridium* sp., *Desulfotomaculum* sp., *Thermincola* sp., *Ruminococcus* sp.), Sphingomonadales and several sulfur reducing bacteria, such as Desulfovibrionales, Desulfuromonadales, and Desulfobacterales were the dominant obligate anaerobic chemoheterotrophic bacteria in this field. The presence and involvement of Clostridiales (Sherry et al., 2013), Sphingomonadales (Leys et al., 2005) and several sulfate reducing bacteria (SRB) (Callaghan et al., 2012b) have been reported in several similar studies.

In addition to the chemoheterotrophic bacteria, there were several members of groups of microorganisms belonging to other metabolic categories. Different types of phototrophic microorganisms were found in the field community with the ability to use H_2O (oxygenic photoautotrophic microorganisms), H_2S (non-oxygenic photoautotrophic bacteria) or an organic compound (photoheterotrophic bacteria) as the source of electrons required for CO_2 fixation process. A portion (4%) of these bacteria belonged to several types of chemolithotrophic bacteria with the ability to use mainly nitrogen, such as Nitrospirales (mainly *Candidatus nitrospira defluvii*) and Nitrosomonadales (*Nitrosomonas* sp.), and iron, especially Acidimicrobiales (*Acidithiomicrobium* sp. P2 and *Acidimicrobium ferrooxidans*). The photoautotrophic bacteria identified in this study fell into two categories; three orders, mostly Chromatiales (1.2%), chlorophyll harbouring-Sphingomonadales (*Erythrobacter* sp.) (0.5%) and a few Chlorobiales (0.1%) belonged to of anoxygenic photoautotrophic bacteria, the other category consisted of different types of oxygenic photoautotrophic bacteria (Cyanobacteria) and eukaryotic green algae. *Erythrobacter* sp. are a group of photoheterotrophic bacteria that obtain their main energy requirements from light, while depending for their growth on organic substrates as carbon sources (Whitman and Parte, 2009, Schink et al., 2011). This genus of bacteria is able to utilise a variety of aromatic and aliphatic hydrocarbons through an aerobic chemoheterotrophic activity (Leys et al., 2005). The presence of both anoxygenic and oxygenic phototrophic microorganisms in the sites contaminated with hydrocarbons, especially crude oil, have been shown in previous studies as well (Jiménez et al., 2007, Raghukumar et al., 2001). The production of molecular oxygen, specifically by the phototrophic microorganisms, as well as nitrogen fixation and the production of simple

organics can be critical for the improvement of bioremediation of hydrocarbons by chemoheterotrophic bacteria (Papizadeh et al., 2011). However, as reported by AN et al (An et al., 2013), the phototrophic microorganisms comprise only a small portion of the microbial communities in the crude oil contaminated sites.

Data analysis using The KEGG (Kyoto Encyclopedia of Genes and Genomes) showed that the metagenome capacity of the crude oil contaminated soil was wide enough to allow most known metabolic pathways required for the growth and activity of microorganisms to proceed. The majority of these genes and metabolic pathways are involved in routine microbial activity, such as intracellular clustering processes, metabolism of carbohydrates, amino acids and proteins, cofactors, vitamins, and belonged mainly to bacteria, especially Proteobacteria and Actinobacteria. However, as we will mention below, a small portion of the genes found in this study are involved in the adaptation of the microbial population to high hydrocarbon concentrations and the degradation of these compounds in order for them to be used as a metabolite or to reduce their toxicity.

3.5.5.1 Alkane monooxygenases

Alkane hydroxylating enzymes are a group of monooxygenases responsible for the oxidation of n-alkanes and several other compounds to their corresponding alkan-1-ol (Li et al., 2008b). Depending on the microorganism involved, and the electron carrier and substrates, these enzymes are classified into alkane rubredoxin-dependent monooxygenase, flavin-binding monooxygenase, CYP153 soluble cytochrome P450 (for C5–C12 alkanes), methane monooxygenase (in methane oxidizing bacteria), propane monooxygenase (in propane-oxidizing bacteria), ammonia monooxygenase (in ammonia-oxidizing bacteria), butane monooxygenases (BMOs) and etc. (Redmond et al., 2010, Sayavedra-Soto et al., 2011). While CYP153 soluble cytochrome P450 monooxygenases are responsible for mono-oxidation of the alkanes with C5–C12 length carbon, flavin-binding monooxygenase (AlmA) oxidize long chain (C20–C36) alkanes. Overall, less than 0.01% of the metagenome belonged to alkane hydroxylating enzymes, mainly categorised as alkane rubredoxin-dependent monooxygenases and Flavin-binding monooxygenases (AlmA; for C20–C36 alkanes). These two genes were found in a variety of phylogenetic groups, including Actinomycetales, Burkholderiales, Pseudomonadales, Oceanospirillales, Deinococcales, Rhodobacterales and Sphingobacteriales (Table 3.5.1). A limited numbers of microorganisms in this study, including *Rhodococcus* sp. H1, *Mycobiacterium marinum* M and *Dietzia cinnamea* P4 (all belonging to Actinomycetales)

and *Salinisphaeraceae bacterium* PC39 (Salinisphaerales) harboured Cyt-P450-dependent monooxygenase genes, as well. Furthermore, the genome of two genera of Actinomycetales, *Mycobacterium* sp. (*M. smegmatis* str. MC2 155, *M. chubuense* NBB4) and *Rhodococcus* sp. (*R. imtechensis, Rhodococcus jostii* RHA1), contained propane monooxygenase. In the studies performed by Liang et al (Liang et al., 2011) on soils contaminated with crude oil, the majority of functional genes involved in hydrocarbon degradation belonged to the alkane-1-monooxygenase of *Rhodococcus erythropolis* and *Rhodococcus* sp. Q15. In addition, the soluble methane monooxygenase (sMM), which catalyses the oxidation of methane, was found in a variety of methanotrophic archaea and several bacteria, including *Acidiphilium* sp., *Amycolatopsis* sp., *Beijerinckia* sp., *Crenothrix* sp., *Frankia* sp., *Marinobacter* sp., *Mycobacterium smegmatis* str. MC2 155, *Verminephrobacter eiseniae* EF01-2 and many ammonia oxidising bacteria (Table 3.5.1). Indeed, methane monooxygenase in methane oxidizing archaea is related to Ammonia monooxygenases (AMOs) found in ammonia oxidising bacteria, which are able to hydroxylate ammonia as well as methane and many other short length alkanes (De La Torre et al., 2008).

3.5.3.2 Genes responsible for degradation of aromatic hydrocarbon

Overall 2.5% of the genes found in the soil samples collected from crude oil well field belonged to functional enzymes responsible for the metabolism of aromatic hydrocarbons (Table 3.5.2). Among these aromatic hydrocarbon degrading genes 0.3% were responsible for initiation of aerobic catabolism of several types of hydrocarbons. Furthermore, the enzymes responsible for initiation of anaerobic aromatic degrading pathways comprised 0.4% of the total number of the functional genes involved in aromatic hydrocarbon degradation. Other genes in this category fall into peripheral pathways for catabolism of aromatic compounds (0.6%) and the genes involved in metabolism of central aromatic intermediates (1.2%). A full list of the genes involved in the degradation of aromatic hydrocarbons found in this study is presented in Table 3.5.2. These genes are potentially able to catabolize a large variety of aromatic hydrocarbons, such as aromatic amines, benzene, toluene, carbazol, naphthalene and Ethylbenzene, through aerobic metabolism. However, the majority of these genes were responsible for encoding Benzoate (0.13%) and Phenylacetate (0.1%) catabolic pathways. Furthermore, approximately 0.4% of the whole metagenome obtained in this study belonged to the genes encoding the enzymes involved in anaerobic degradation of aromatic hydrocarbons, mostly benzoate.

Despite difference in the ratios of these genes and their sequences, a similar list of genes responsible for the degradation of aromatic hydrocarbons was reported in previous studies (Liang et al., 2011, An et al., 2013) working on metagenomic studies of crude oil contaminated soils.

Enzymes		Orders	Ratios
Aliphatics	Alkane rubredoxin-dependent monooxygenase	Actinomycetales, Burkholderiales, Pseudomonadales, Flavobacteriales, Legionellales, Oceanospirillales, Rhodobacterales, Rhodospirillales, Myxococcales, Cytophagales, Deinococcales,, Solirubrobacterales, Sphingobacteriales	0.006%
	Alpha-ketoglutarate-dependent dioxygenase AlkB	Burkholderiales, Enterobacteriales	0.00004%
	Bacterial Flavin-binding monooxygenase (AlmA; for C20–C36 alkanes)	Actinomycetales, Alteromonadales, Burkholderiales, Rhodobacterales, Pseudomonadales, Oceanospirillales, Rhizobiales, Deinococcales, Rhodobacterales, Sphingobacteriales, several Fungi *(Ajellomyces dermatitidis* SLH14081, *Coccidioides posadasii* str., *Grosmannia clavigera* kw1407, *Metarhizium anisopliae* ARSEF 23)	0.003%
	Propane monooxygenase	Actinomycetale	0.0005%
	Bacterial CYP153: cytochrome P450 (for C5–C12 alkanes)	Actinomycetales, Salinisphaerales	0.0005%
	Methane monooxygenase/ammonia monooxygenase, subunits A, B and C	Actinomycetales, Methylococcales, Nitrosomonadale, Nitrosopumilales, Alteromonadales, Burkholderiales, *Chromatiales*, Rhizobiales, Rhodospirillales, Sphingobacteriales	0.0008%
Aromatics	Aerobic pathways	Aromatic Amine Catabolism, Benzoate transport and degradation cluster, Cresol degradation, Gentisarate degradation, Phenylacetyl-CoA catabolic pathway (core), Toluene 4-monooxygenase (T4MO), carbazol degradation cluster p-cymene degradation	0.3%
	Anaerobic pathways	Acetophenone carboxylase 1, Anaerobic benzoate metabolism, Anaerobic toluene and ethylbenzene degradation	0.4%
	Central aromatic intermediates	4-Hydroxyphenylacetic acid catabolic pathway, Catechol branch of beta-ketoadipate pathway, Central meta-cleavage pathway of aromatic compound degradation, Homogentisate pathway of aromatic compound degradation, N-heterocyclic aromatic compound degradation, Protocatechuate branch of beta-ketoadipate pathway, Salicylate and gentisate catabolism	0.6%
	Peripheral pathways	Benzoate catabolism, Biphenyl Degradation, Chloroaromatic degradation pathway, Naphtalene and antracene degradation, Phenol hydroxylase, Phenylpropanoid compound degradation, Quinate degradation, Salicylate ester degradation, Toluene degradation, n-Phenylalkanoic acid degradation, p-Hydroxybenzoate degradation	1.2%

Table 3.5.1. List of the genes and microorganisms involved in degradation of aliphatic and aromatic hydrocarbons identified in the metagenomic data analysis crude oil filed soil.

3.5.3.3 Genes responsible for desulfurization of sulfur-containing hydrocarbon

The desulfurization of sulfur-containing hydrocarbons is one of most important points of interest due to their role in providing sulfur supplementation for hydrocarbon remediating organisms and also in improving the quality of oil-originated fuels (Milam et al., 2014). Several microorganisms have been found with the hydrocarbon desulfurization ability, which employ different enzymatic reactions to remove these elements from their substrates (Firouz Abbasian et al., 2015). These enzymes can be categorized into Alkanesulfonates monooxygenase, which remove sulfur from sulfonated alkanes, and aromaticsulfonate desulfurating enzymes such as Dibenzothiophene (DBT) degrading enzymes, which convert a poly-aromatic sulfur heterocyclic compound to 2-hydroxybiphenyl (2-HBP) (Firouz Abbasian et al., 2015). The later enzyme is ordered usually as an operon consisted of three genes, including dszC (dibenzothiophene monooxygenase), dszA (Dibenzothiophene sulfone monooxygenase) and dszB (2,2-hydroxybiphenyl benzensulfunate desulfinase). The genes encoding both groups of alkane and DBT desulfurating enzymes were found in these metagenomic studies. The genes responsible for desulfurization of Alkanesulfonate, including Alkanesulfonate monooxygenase and Alkanesulfonate transport system, consisted 0.024% of whole microbial genome in these samples, mostly carried by different species of Actinomycetales, Clostridiales, Burkholderiales, Pseudomonadales and Rhizobiales. The genes responsible for desulfurization of dibenzothiophene comprised a minority of the hydrocarbon desulfurizing genes, and were mostly carried on Actinomycetales, Bacillales, Burkholderiales and Rhizobiales. The DNA sequences of many of the Alkanesulfonate and DBT desulfurizing genes found in this study showed significant differences with the existing corresponding genes in public gene databases. Since the DNA sequences of genes give no idea of their expression and enzymatic activities, further studies are required to isolate and express these genes in laboratories.

Pathway	Genes	Microorganisms	Ratio
Alkanosulfate degrading pathway	Alkanesulfonate monooxygenase (EC 1.14.14.5)	Actinomycetales, *Bacillales*, Burkholderiales, Enterobacteriales, Pseudomonadales, Rhizobiales, Caulobacterales, Herpetosiphonales, Rhodospirillales, Sphingomonadales, several Algae and Cyanobacteria such as *Anabaena variabilis* ATCC 29413, *Nostoc punctiforme* PCC 73102, *Microcystis aeruginosa* NIES-843,	0.004%
	Alkanesulfonates transport system permease protein	Actinomycetales, Burkholderiales, Clostridiales, Enterobacteriales, Pseudomonadales, Rhizobiales, Rhodospirillales, Rhodobacterales, Desulfuromonadales, Bifidobacteriales, Thermoanaerobacterales, Sphingomonadales, Solirubrobacterales, Methanosarcinales, Methylophilales, Synergistales, an unclassified bacterium (*Thermobaculum terrenum* ATCC BAA-798), Caulobacterales and Myxococcales	0.02%
Dibenzothiophene degrading enzymes	dibenzothiophene (DBT) monooxygenase (dszC)	Actinomycetale, *Bacillales*, Burkholderiales, Rhizobiales and Rhodospirillales	0.002%
	DBT-sulfone monooxygenase (dszA)	Actinomycetales, *Bacillales*, Burkholderiales, Pseudomonadales and Rhizobiales	0.001%
	dibenzothiophene-5,5-dioxide monooxygenase (dszB)	Actinomycetales, *Bacillales*, Burkholderiales, Pseudomonadales, Rhizobiales, Solirubrobacterales and *Cyanobacteria*	0.001%

Table 3.5.2. List of the genes and microorganisms involved in desulfurization of hydrocarbons identified in crude oil metagenomic data analysis.

3.5.3.4 Genes responsible for microbial adaptation and growth in hydrocarbon contaminated soils

Biosurfactants are a group of amphiphilic surface active agents with the ability to reduce surface/interfacial tensions between two immiscible fluids (Saharan BS et al., 2011). Due to improvement in the solubility of polar compounds in organic solvents, biosurfactants are used in several crude oil dependent technologies, such as transfer of crude oil, enhanced oil recovery and crude oil bioremediation (Saharan BS et al., 2011). Due to the insolubility of hydrocarbons in water, the production of these biosurfactants looks likely to be necessary for improvements to the bioavailability of these compounds (Soberón-Chávez and Maier, 2011). The majority (0.8%) of biosurfactant-encoding genes in this study produce trehalose, which is a disaccharide residue on a glycolipid surfactant substituted by mycolic acids at C-6 and C-6′ positions (Lang et al., 2000). Polyol lipids, including (non)oxyethylenated ethers or esters of a polyol, were the second dominant (0.015%) biosurfactant producing genes. In addition, bacterial exopolysaccharide mono/di-rhamnolipids, which are produced by the action of

rhamnosyltransferase 1 and rhamnosyltransferase 1 (Müller et al., 2012), were the third most dominant (0.001%) group of biosurfactants in this study.

Macro/micro-elements are essential for normal growth of microorganisms, and any shortage of these elements restrict microbial growth. Depending on the elements and the enzymatic capacity of microorganisms, these essential nutrients are supplied from mineral or organic compounds present in the microbial environment. Therefore, microorganisms have to spare a portion of their genome capacity for improving the bioavailability of different elements. Since crude oil is a very deprived environment for many essential elements, especially phosphorus, nitrogen, sulfur, any contamination can lead to changes in the C:N:S:P ratio (Uhlik et al., 2013). This is the scenario expected in the soil samples obtained from the crude oil contaminated field where the microorganisms spare a significant portion of their genome capacity for improving the bioavailability and acquisition of phosphorus, nitrogen, sulfur, potassium and iron (1.63%, 1.53%, 1.34%, 0.4%, 0.7% respectively) (Table 3.5.3).

Furthermore, since crude oils contain different heavy elements, such as arsenic, cadmium, cobalt, zinc, cadmium, copper, mercury, chromium, zinc and many other minerals, these may have toxic effects on microbial growth and activities, (Osuji and Onojake, 2004). Therefore, the genomes of the microorganisms present in these environments must contain the genetic capacity to enable microorganisms to be resistant to these toxic elements. Genes resistant to several types of toxic elements were found in the dataset obtained from our given sample (Table 3.5.3). The genes responsible for microbial resistance to heavy metals found in this study can be classified into two categories. The first group of these genes express several cell membrane integrated channels, such as Co/Zn/Cd efflux system membrane fusion protein, cadmium-transporting ATPase, *copA* ATPase dependent transporter, cusCFBA proton-cation antiporter complexes, mercuric transport protein and arsenic efflux pumps, which are involved in excretion of the heavy metals. The second category of these genes, such as multicopper oxidase, organomercurial lyase, mercuric reductase and arsenate reductase, are responsible for reducing the toxicity of the heavy metals through their oxidation, reduction or conjugation with a (in)organic compound.

Ability	Elements	Mechanism	Ratios
Metal resistance genes	Arsenic	Arsenate reductase, Arsenic efflux pump protein, Arsenic resistance protein ArsH, Arsenical pump-driving ATPase, Arsenical-resistance protein ACR3, Arsenical regulatory operons	0.1%
	Cadmium	Cadmium resistance protein, Cadmium-transporting ATPase	
	Cobalt-zinc-cadmium	Cadmium-transporting ATPase, Cation efflux system proteins CusA, Cobalt-zinc-cadmium resistance proteins CzcA and CzcD, Heavy metal RND efflux outer membrane protein (CzcC family), Copper sensory histidine kinase CusS, Heavy metal sensor histidine kinase, Hypothetical protein involved in heavy metal export, Nickel-cobalt-cadmium resistance protein NccB, Putative silver efflux pump, Transcriptional regulatory genes (MerR family, Cd(II)/Pb(II), CusC, CusR, HmrR,	0.007%
	Copper	CopG protein, Copper chaperone, Copper resistance proteins C and D, Copper-translocating P-type ATPase (EC 3.6.3.4), Multicopper oxidase, Magnesium and cobalt efflux protein CorC, CutF precursor, CusS, CusR,	0.84%
	Mercury	Mercuric ion reductase, Mercuric transport proteins (MerC, MerE and MerT), Organomercurial lyase, Periplasmic mercury(+2) binding protein and Mercuric resistance operon regulators	0.09%
	Chromium	Chromate resistance proteins: ChrB, ChrA, ChrC Superoxide dismutase SodM-like protein ChrF	0.03%
	Zinc	Response regulator of zinc sigma-54-dependent two-component system, Zinc resistance-associated protein	0.02%
Genes involved in element shortage	Sulfur	Arylsulfatase, Sulfite reductase (HmeA, HmeD, HmeC, HmeB), Sulfur oxidation (Sulfide dehydrogenase, Sulfite oxidase, SoxC, SoxB, SoxD, SoxY, SoxZ, SoxA, SoxR, SoxS, SoxH), Thioredoxin-disulfide reductase, Inorganic Sulfur Assimilation (sulfur ABC-type transporter, Adenylyl-sulfate reductase, Adenylylsulfate kinase, Adenylylsulfate reductase), uptake and metabolism of Alkanesulfonate, DMSP, L-Cystine, Taurine and glutathione	1.34%
	Nitrogen	Allantoin Utilization, Amidase, Ammonia assimilation (Glutamate synthase, Glutamate-ammonia-ligase adenylyltransferase, Glutamine amidotransferase), Cyanate hydrolysis, Denitrification (nitrite reductase, Nitric oxide reductase, Nitrous oxide reductase and their accessory proteins), nitrite reductase, Nitric oxide synthase oxygenase, Nitric-oxide synthase, Nitrilase, Nitrogen fixation (Nitrogenase and its associated proteins)	1.53%
	Phosphorus	Alkylphosphonate utilization, Phosphate ABC transporter, Phosphate-binding DING proteins	1.63%
	Iron	Ferrous and Ferric Transport system, Heme uptake and utilization, Iron(III) dicitrate transport system Fec, Siderophores (Alcaligin, Achromobactin, Aerobactin, Anthrachelin, Bacillibactin, Enterobactin, Petrobactin, Pyoverdine, Petrobactin, Salmochelin, Desferrioxamine E, pyochelin, Staphylobactin, Yersiniabactin, Vibrioferrin)	0.7%
	potassium	efflux system ATP-binding protein and ancillary proteins (KefG, and KefC), K transporter trk, KdpE, KdpD, K/H antipoter ROSB, Kef-type transport system 2, FKBP-type peptidyl-prolyl cis-trans isomerase FkpA precursor, Glutathione-regulated potassium-efflux system ATP-binding protein, Kup system potassium uptake protein, Large-conductance mechanosensitive channel, Potassium channel protein, Potassium voltage-gated channel subfamily KQT, Potassium-transporting ATPase A, B and C chains, Voltage-gated potassium channel, cAMP-dependent Kef-type K+ transport system, pH adaptation potassium efflux system A	0.4%

Table 3.5.3. Ratios of the genes involved in metal resistance and shortage of elements identified in crude oil metagenomic data.

3.5.6 Conclusion

To sum up, based on this metagenomic study, the soils contaminated with crude oil in the oil field area were inhabited by a diversity of microorganisms in which Actinomycetales, Rhizobiales, Burkholderiales and Clostridiales were found as dominant microorganisms. These microorganisms are rich sources of different enzymes useful for the degradation of aliphatic and aromatic hydrocarbons, the production of biosurfactants, and the improvement of the quality of petroleum throughout crude oil refining process. Furthermore, the genetic capacity of these microorganisms provides valuable information regarding the adaptation of hydrocarbon degrading microorganisms to different harsh environmental conditions, such as heavy metal toxicity, hydrocarbon hydrophobicity and the C:N:P imbalances, which are useful for the design of bioremediation systems for hydrocarbon contaminated sites. However, since crude oil wells have been established in different climatic conditions, from hot deserts to frozen tundra, the need for further studies is required to obtain more detailed information regarding the presence and participation of different microbial strains and their corresponding hydrocarbon degrading genes and to assess the activity of each stain in these differing conditions.

3.6 Microbial diversity and functional gene capacity of microorganisms in crude oil

3.6.1 Abstract

While crude oil is still considered likely to be the main source of energy in this century, oil spillages can be disastrous for the environment. Crude oil originating from a particular area also has a specific fingerprint which can affect its value. In this study, the metagenome of a crude oil environment was investigated using metagenomics sequencing and the MG-RAST programmes to identify the components of the microbial community and their functional gene capacity. Several genes coding for enzymes degrading aliphatic and aromatic compounds were found in this study, which are potentially useful for the removal of crude oil contamination through bioremediation. Furthermore, a list of genes responsible for producing desulfurase enzymes were identified which can be used for the improvement of the quality of crude oil.

3.6.2 Introduction

Although there are many limiting factors for microbial growth in crude oil, especially the imbalance in the ratios of carbon sources to those of nitrogen and phosphate (C:N:P), it is inhabited by a variety of microbial populations that can utilise hydrocarbons as a source of carbon and energy (Abbasian et al., 2015a). Since crude oil microorganisms metabolise these hydrocarbons, they can positively or negatively affect the quality of crude oil (Abbasian et al., 2015a). In the Microbially Enhanced Oil Recovery (MEOR) technique , microorganisms are used to improve the quality of crude oil through the production of different compounds such as surfactants, gases, acids, alcohols and polymers (Siegert et al., 2014). Inversely, since microorganisms metabolise the hydrocarbons present in crude oil, they can also damage the quality of crude oil as a result of changes in acidity and viscosity (Berdugo-Clavijo and Gieg, 2014, Larter and Head, 2014). In addition , since these microorganisms are likely to be rich source of the enzymes responsible for the degradation of hydrocarbons, they can be potentially useful for the remediation of hydrocarbon and crude oil spillages (Mail et al., 2012, Abbasian et al., 2016b).

Despite broad investigations on the microbial flora in crude oil, there is not enough knowledge regarding the microbial diversity present in crude oil and the genes involved in occupying that environmental niche. Metagenomic techniques based on the random sequencing of environmental DNA and bioinformatic analysis of the data provides significant information regarding the microbial diversity, genome capacity, function of the genes and the metabolic

pathways active in a given environment (Abbasian et al., 2015b). Although PCR amplification has been used in the past for the identification of genes present in a DNA sample, in a metagenomic study the DNA is directly sequenced and thus it is theoretically possible to determine the sequence of genes and their function without any requirement for prior knowledge of the gene sequence (Gong et al., 2013). This approach is potentially useful for studying the microbial population in complex environmental samples, including crude oil, in which many of the microbial flora cannot grow in standard microbial media due to special requirements for unknown physicochemical factors (Abbasian et al., 2015b). This information helps us to identify the active members of an oil-utilising community and its probable changes in response to environmental parameters and the presence of specific oil substrates (Cecchini et al., 2013, Gong et al., 2013). This information is also useful for the development of optimized chemical, physical and environmental conditions that may improve biodegradation of oil and its products (Cecchini et al., 2013, Gong et al., 2013). Furthermore, this approach can also assist in the establishment of correlations between oil components and microbial diversity (Head et al., 2006). To achieve these aims, a metagenomic approach using the Illumina HiSeq platform was used in this study to investigate microbial diversity, genome capacity and functional analysis of organisms present in crude oil.

3.6.3 Material and methods

Preparation of crude oil samples. The crude oil samples were prepared from crude oil pipelines located in oil fields, Perth, Australia. The samples were transferred to the microbiology laboratory under cold conditions.

3.6.3.1 TPH analysis

The sequential ultrasonic solvent extraction method (Risdon et al., 2008) was used to extract TPH compounds from soil sediment samples. Quantification of TPH concentration was carried out using a gas chromatograph fitted with a flame ionisation detector (GC-FID, Agilent 6890). Chromatography was performed on a fused-silica capillary column BPX-5 from SGE (15 m × 0.32 mm internal diameter) coated with HP-5 (0.10-μm film thickness). Helium was used as the carrier gas at 2.5 mL min^{-1}, and the FID detector temperature maintained at 300°C. Splitless injection with a sample volume of 1 μL was applied. The oven temperature was increased from 50 °C to 300 °C at a gradient of 25 °C min^{-1} and held at this temperature for 5 min. The total run time was 19.6 min. Hydrocarbons were quantified using Agilent Chemstation Software by

integration and calibration of peaks of a known concentration of an external calibration standard– Hydrocarbon Window Defining Standard (C_8-C_{40}) from AccuStandard®(Miller and Miller, 1987). Five concentrations of external calibration standard in the range expected in the samples were analysed; a linear curve fit with a R^2 value of 0.997 was obtained. The CCV (Continuing Calibration Verification) was analysed at the start and end of every 20 samples and CCV recovery was 95-110% of true value. Hexane was run as blank with every 10 samples to demonstrate that the system was free from contamination. The surrogate (*o*-terphenyl) was spiked at a level to produce a recommended extract concentration of 20 µg/mL. Surrogate recoveries lay in the range 70-120% for all the samples analysed. 35 mg kg^{-1} was the minimum concentration of TPH detected (MDL) through this analytical method.

3.6.3.2 DNA extraction and DNA sequencing using the Illumina Hiseq platform

Overall, two 50 ml crude oil samples were centrifuged at 5000 RPM for 10 minute and the pellets were used for DNA extraction. DNA extraction and DNA quantification were performed using the power DNA kit (MO BIO) and quantifluor dsDNA system (Promega), respectively, both according to manufacturer's instructions. The DNA samples were sequenced on the Illumina Hiseq platform provided at AGRF (Australian Genome Research Facilities).

3.6.3.3 Data analysis

The raw data obtained from pyrosequencing of the metagenome were trimmed by QIIME (Quantitative Insight Into Microbial Ecology) in order to eliminate uninformative and/or duplicated sequences. Following submission of the reads in Metagenome Rapid Annotation using Subsystem Technology (MG-RAST) (Glass and Meyer, 2011), the data were analysed by the M5RNA annotation source (a minimum alignment length cutoff of 50 bp, an E-value cutoff of 1e-5 and a minimum percentage identity cut off of 97%) to analyse the microbial diversity in this sample. The best hit classification of genes was determined using M5NR annotation source and a minimum alignment length cutoff of 15 bp, an E-value cutoff of 1e-5 and a minimum percentage identity cut off of 70%. Furthermore, the functional abundances of the genes and their hierarchical classification were analysed through COG (cluster of orthologous groups), NOG (non-supervised orthologous groups) and KO (KEGG orthology) subsystem sources using cutoff of 15 bp, an E-value cutoff of 1e-5 and a minimum percentage

identity cutoff of 70%. In addition to the analysis programme provided by MG-RAST, the data were transferred to excel software for further analysis.

3.6.3.4 Nucleotide sequence accession numbers

The metagenomic analysis of crude oil are publicly available in the MG-RAST system as A11, A12, A31 and A32 under project identifiers 4583536.3, 4583535.3, 4583734.3 and 4583735.3, respectively.

3.6.5 Results and discussion

Metagenomics is a powerful tool for the investigation of microbial communities using the genetic information of the DNA obtained from environmental samples (Abbasian et al., 2015b). This approach is useful for the study of the functional abilities and metabolic capabilities of the strains identified in the given habitat, as well (Abbasian et al., 2015b). Using this technique it is possible to investigate microbial communities and their niche in an environment without any requirement to be able to culture the microorganisms in media (Tringe et al., 2005). Furthermore, this technique is quite cost effective in comparison to other molecular approaches and the biases caused by PCR and cloning are not seen in metagenomic sequencing (Abbasian and Saberbaghi, 2013a, Wilson and Piel, 2013). In this study, a HiSeq Illumina platform was used to analyse the microbiome of crude oil obtained from oil wells located in Perth, Australia.

3.6.5.1 Gas chromatograph assays

Based on the gas chromatography assays, the average levels of hydrocarbons (mostly>C10-C16, >C16-C34 and >C34-C40) in crude oil samples were measured as (18421, 33203 and 30 ppm, respectively).

3.6.5.2 Statistics of sequences produced by HiSeq Illumina

Sequencing of these two DNA samples obtained from microbiome of crude oil by HiSeq Illumina platform produced four metagenomics data from each samples (A11, A12, A21 and 22), which were uploaded on MG-RAST for bioinformatic analysis. The statistics of sequences produced by HiSeq Illumina are given in Table 3.6.1.

Sample	A11 and A12	A31 and 32
Total bp count	2,460,382,350	3,075,200,700
Total Sequences count	16,402,549	20,501,338
Post QC: Sequences Count	14,160,659	18,814,626
Average length (bp)	150 ± 0 bp	150 ± 0
Post QC: Mean Sequence Length (bp)	143 ± 20	142 ± 21
Predicted rRNA features	146,442	181,111
Identified rRNA features	2,387	3,047
Predicted known functional genes	7,294,896	12,005,454
Identified unknown functional genes	2,775,419	3,969,907
GC%	46 ± 13 %	48 ± 11 %
H index	3.5	3.3
D index	33.0	25.5

Table 3.6.1. Total data achieved by pyrosequencing process on crude oil, oil tanker sludge and oil contaminated soil.

3.6.5.3 Phylogenetic compositions of the bacterial communities

These datasets were analysed in MG-RAST against M5RNA annotation source using maximum e-value cutoff of 1e-5, minimum identity percentage cutoff of 97% and minimum alignment length cutoff of 50 bp. Phylogenetic trees of the organisms obtained from the samples are shown in Figure 3.6.1. As it can be seen in the Figure 3.6.2, the majority of the organisms (79%) in the sample belonged to bacteria, followed by archaea (12%) and Eukaryota (3%). In total 6% of the sequences could not be classified at the phylum level based on current databases in RDP annotation source (Fig. 3.6.2). These unclassified sequences were submitted to NCBI as uncultured organisms obtained from crude oil (accession numbers KM872737-KM873025). Overall, crude oil was inhabited by Proteobacteria (21.4%), Bacteroidetes (12%), Euryarchaeota (11.5%), Firmicutes (4.2%), Spirochaetes (2.4%) and many other phyla with

lower than 2% abundance (Fig. 3.6.2). Although the proportion of microorganisms is different in this study from that obtained by Silva and colleagues (Silva et al., 2013), most of dominant microorganisms listed here were also identified in petroleum samples from Brazilian oil fields. Furthermore, a similar study performed on the soils collected from the crude oil wells located at Perth, the same geographical condition as this current research, showed that contamination with soil can change the ratios of microorganisms, leading to increases in the numbers of Actinobacteria, Firmicutes, Acidobacteria, Planctomycetes, Gemmatimonadetes and many other bacterial phyla, and decreases mainly in the ratios of Bacteroidetes, Deferribacteria and Chloroflexi (Firouz Abbasian et al., 2016). These changes depends on the geochemistry of the soil, and can be variable in different soil sources. For instance, as it was shown before in an experimental soil collected from a farm at Adelaide, South Australia, contaminated with the same crude oil worked on this experiment (Abbasian et al., 2016a). In this study, Actinomycetales, Flavobacteriales, Glomerellales, Planctomycetales, Verrucomicrobiales, Rhizobiales, Rhodospirillaes, Cytophagales and Bacillales were the most dominant microorganisms.

Furthermore, regardless of the Eukaryota and unclassified organisms (42%), classification of the microorganisms at order levels showed that Methanocellales (mostly *Methanocella paludicola* SANAE) (9.8%) was the dominant order present in the crude oil sample, followed by Bacteroidales (6.1%) (*Bacteroides* sp., *Porphyromonas* sp. and *Parabacteroides* sp.,), Syntrophobacterales (*Desulforhabdus* sp. DDT) (3.8%), Burkholderiales (3.5%) (mostly *Alcaligenes faecalis* subsp. Faecalis), Flavobacteriales (mainly *Flavobacterium* sp. and *Capnocytophaga* sp.) (3%), Spirochaetales (*Spirochaeta* sp.) (2.8%), Desulfuromonadales (2.3%) (mostly *Geobacter* sp., *Desulfuromonas* sp. and *Pelobacter* sp.), Clostridiales (Symbiobacterium sp., Desulfotomaculum sp. and Clostridium sp.) (2%), Sphingobacteriales (*Pedobacter* sp., *Sphingobacterium* sp. *Chitinophaga* sp. and *Terrimonas* sp) (2%) and many other orders with less than 2% ratios (Fig. 3.6.2). While more than half portion of the reads in this study belonged to aerobic heterotrophic bacteria, the dominant microorganisms belonged to anaerobic bacteria, including methanogenic archaea (13%), sulfate reducing bacteria (SRB) (4.0%), Bacteroidetes (12%), Syntrophobacterales (3.9%), Spirochaetales (2.8%) and Clostridiales (2%). The presence of the most dominant methanotrophic microorganisms in this study, *Methanocella paludicola* SANAE, and their involvement in degradation of hydrocarbons have been shown in different studies (Chunshuang et al., 2014, Zeman, 2012). Furthermore, there are several reports of involvement of different anaerobic bacteria, such as

Sphingomonadales (Whitman and Parte, 2009, Leys et al., 2005), Bacteroidetes (Whitman and Parte, 2009) and SRBs (Callaghan et al., 2012b, Neria-González et al., 2006, Muratova et al., 2005, Sturchio et al., 2007, Widdel et al., 2010a) in anaerobic biodegradation of aliphatic and aromatic hydrocarbon compounds. Bacteroidales are a common order of microorganisms found in crude oil and crude oil contaminated sites (Wang et al., 2011, Wang et al., 2014b, Gieg, 2010). In addition to these obligatory anaerobic bacteria, several facultative anaerobic microorganisms, including Sphingobacteriales (2.0%), Actinomycetales (0.9%), Bacillales (0.9%), Lactobacillales (0.9%), Rhodocyclales (0.3%) and many others were detected in these crude oil samples. While the ratios of Actinobacteria and Firmicutes in the crude oil samples were much lower than that in comparison with the metagenomics studies performed on crude oil contaminated soils (Firouz Abbasian et al., 2016, Yergeau et al., 2012a, Liang et al., 2011), several reports have shown the involvement of these groups of facultative or obligatory anaerobic Gram positive bacteria in the degradation of hydrocarbons (Sherry et al., 2013).

The rest of the dominant microorganisms in this study either have been found in hydrocarbon contaminated sites or are actively involved in degradation of these compounds. Burkholderiales and Pseudomonadales, for instance, are two very common orders of microorganisms which have been isolated from oil contaminated sources and are involved in aerobic degradation of both aromatic and aliphatic hydrocarbons (Akbari and Ghoshal, 2015, Jayamani and Cupples, 2015, Engel and Gupta, 2014, Wang et al., 2014b, Al-Mailem et al., 2015).

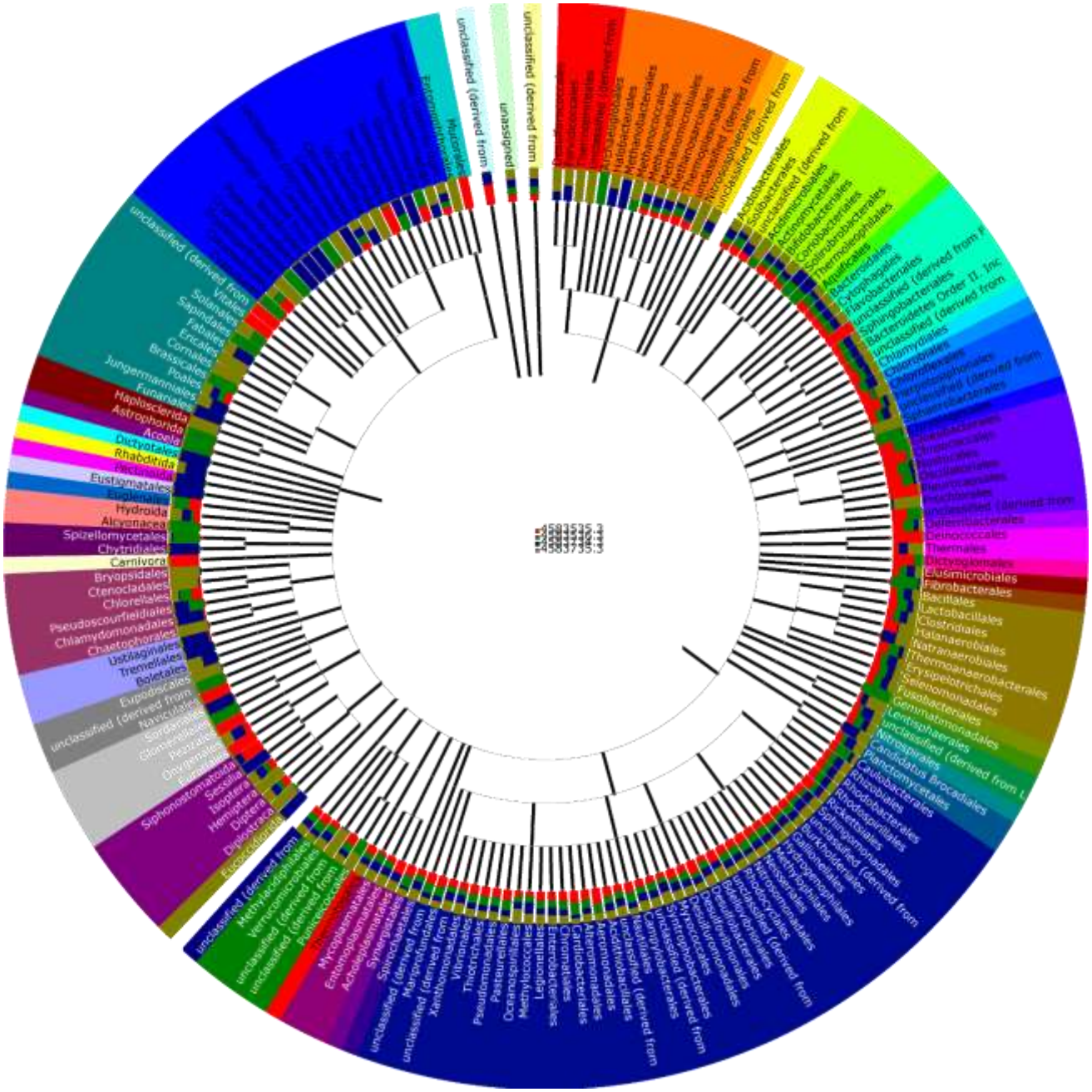


Figure 3.6.1. The microbial orders detected in crude oil samples based on metagenomic analysis.

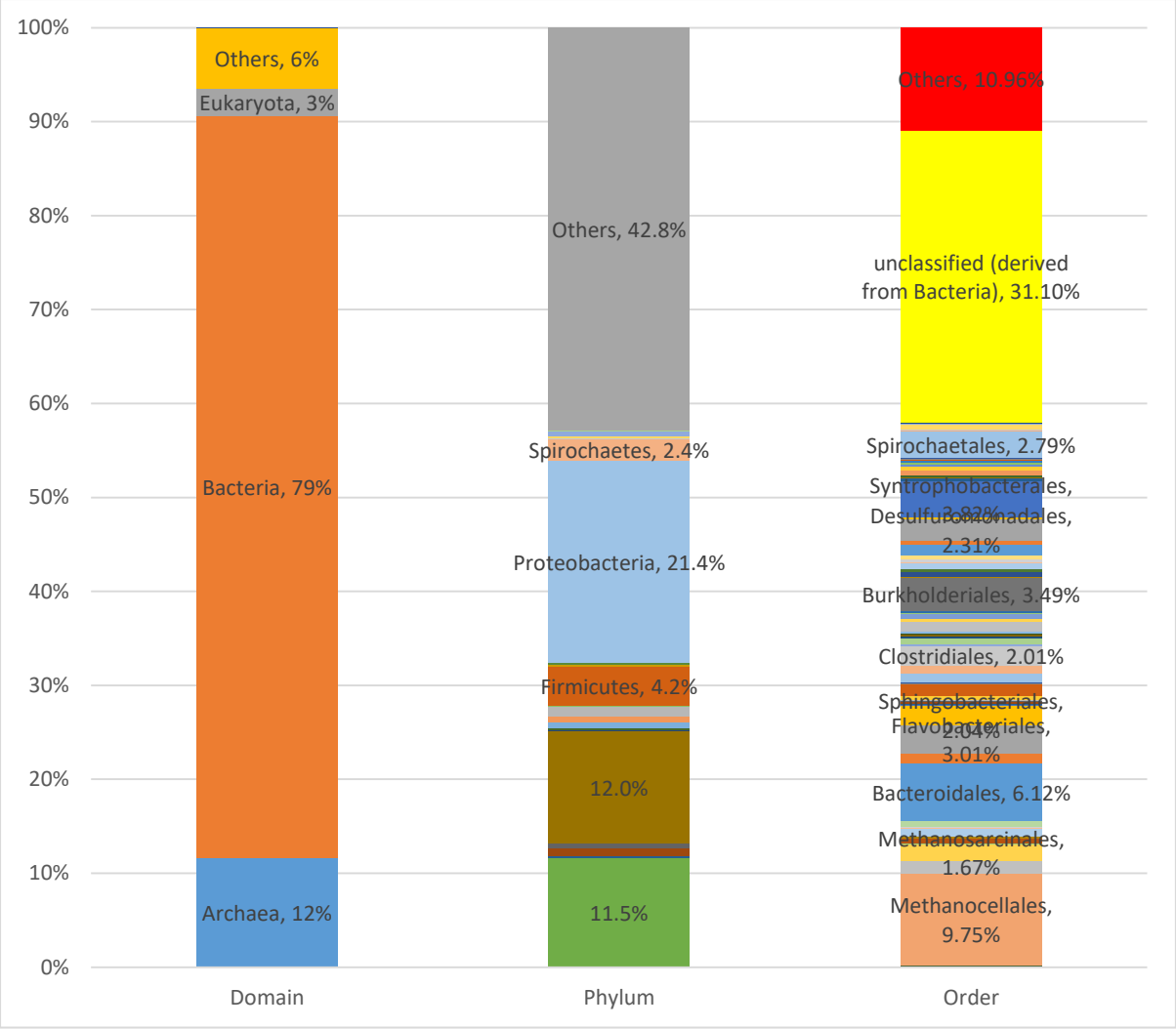


Figure 3.6.2. The ratios of microbial diversity (at domain, phylum and order levels). The name of microorganisms with lower than 1.5% are not mentioned here.

3.6.5.4 Global gene expression and metabolic potential of the crude oil microbiome

To identify the best fit classification of the functional genes in the sample, the database was analysed against M5NR annotation source using maximum e-value cut-off of 1e-5, minimum identity percentage cut-off of 70% and minimum alignment length cut-off of 15 bp. This minimum identity percentage cut-off was chosen since the list of microorganisms in this identity filtration showed more overlaps with the list of microorganisms achieved from phylogenetic composition. This analysis showed that 83.1% of the functional genes present in the microbial community of crude oil belonged to bacteria, followed by archaea (2.7%) and Eukaryota (0.9%). In total 13.5% of the sequences could not be classified at the phylum level based on current databases in RDP annotation source (Fig. 3.6.3). The majority (50.5%) of

these functional genes belonged to Proteobacteria, followed by Bacteroidetes (7.8%), Firmicutes (6.6%), Deferribacteres (5.0%), Actinobacteria (2.8%), Euryarchaeota (2.3%) and many other phyla with lower than 2% abundance (Fig. 3.6.3). Furthermore, regardless of the Eukaryota and unclassified organisms (14.5%), classification of the microorganisms at order levels showed that Burkholderiales (12%) (mostly *Alcaligenes faecalis* sub sp. Faecalis) was the dominant order present in the crude oil sample, followed by Deferribacteriales (5.1%) (mostly *Deferribacter desulfuricans* SSM1, *Denitrovibrio acetiphilus* DSM 12809), Bacteroidales (4.4%) (*Bacteroides* sp., *Porphyromonas* sp. and *Parabacteroides* sp.,), Desulfuromonadales (4.0%) (mostly *Geobacter* sp., *Desulfuromonas* sp. and *Pelobacter* sp.), Rhizobiales (3.1%) (mostly *Candidatus Liberibacter* solanacearum, *Xanthobacter* sp., *Rhizobium* sp. and *Bartonella* sp.), Chromatiales (3%) (mostly *Nitrosococcus oceani*), Rhodocyclales (2.9%) (mostly *Dechloromonas* sp., *Thauera sp.*, *Aromatoleum* sp. and *Azoarcus* sp.), Pseudomonadales (2.5%) (mostly *Psychrobacter* sp.), Actinomycetales (2.5%) (mostly *Mycobacterium* sp. and *Propionibacterium* sp.) and many other orders with less than 2% ratios. This composition of microorganisms carrying the functional genes was different from the data reported before for crude oil contaminated soil in which, regardless of Proteobacteria, the majority of functional genes in this environment were carried by Actinobacteria, Planctomycetes, Acidobacteria, Bacteroidetes and Firmicutes (Firouz Abbasian et al., 2016). However the physico-chemical conditions governing an environment can also affect the microbial diversity, and therefore the functional gene composition, of an environment.

Submission of the metagenome data to KEGG (Kyoto Encyclopedia of Genes and Genomes) to investigate the complete metabolic pathways supporting microbial growth in crude oil showed that the whole microbial metagenome covers all the metabolic pathways required for microbial growth. These protein encoding genes are involved in a range of microbial metabolic systems of which most were involved in cellular activities such as clustering (15%), protein metabolism (10%), carbohydrate metabolism (9.5%), amino acid metabolism (9.0%), miscellaneous (7.5%), cofactors and vitamins (6.5%), DNA metabolism (5%) with many others housekeeping genes at less than 5% (Fig. 3.6.4), which did not show big differences with the same data obtained from crude oil contaminated soil performed by Abbasian et al (Firouz Abbasian et al., 2016). However, since this study focussed on the genes involved in crude oil metabolism, these metabolic groups were not examined further. Through metabolism of crude

oil compounds, microorganisms must be able to tolerate the harsh conditions in that environment and degrade both aliphatic and aromatic hydrocarbons.

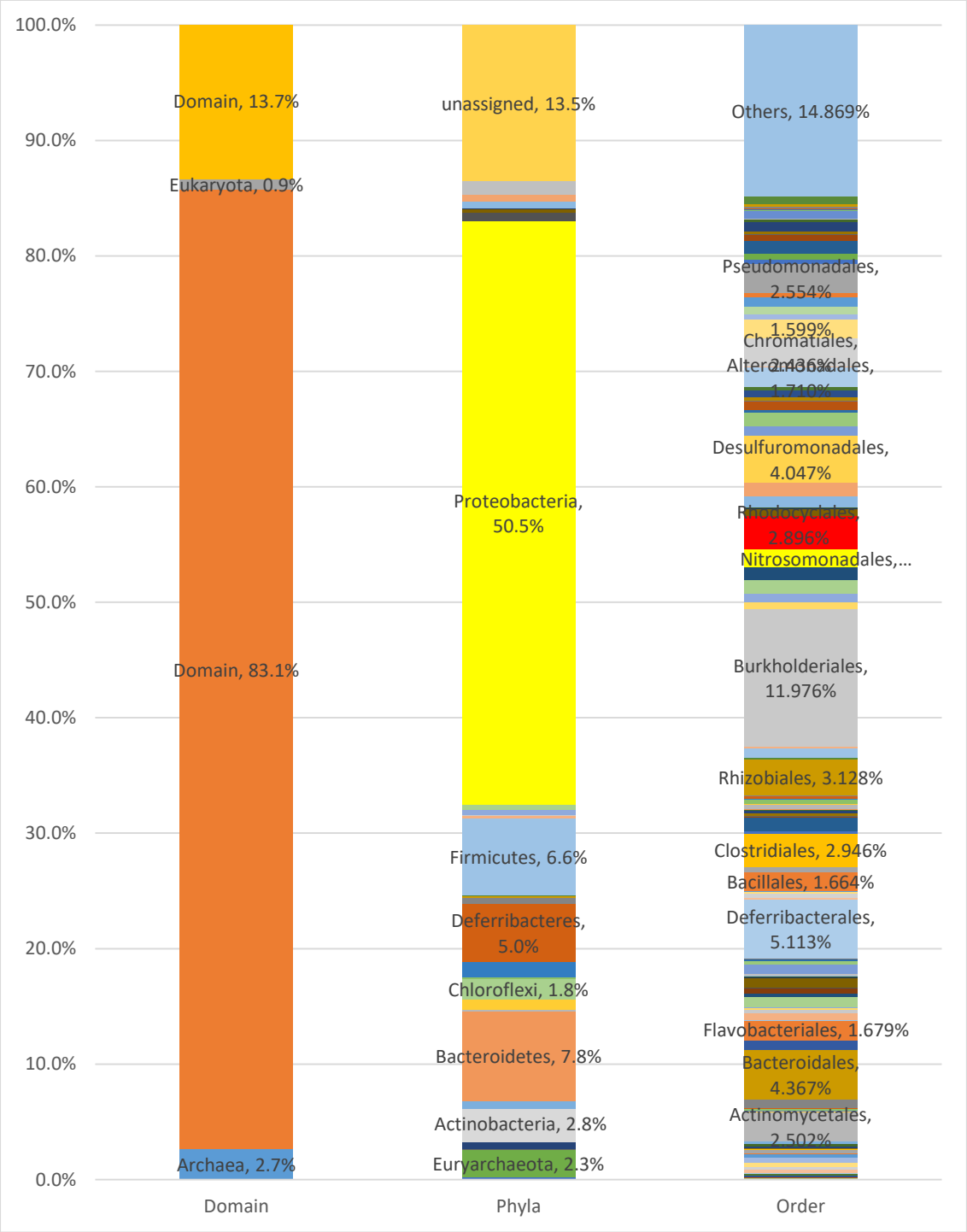


Figure 3.6.3. The ratios of microbial diversity (at domain, phylum and order levels) classified by functional gene capacity. The name of microorganisms with lower than 1.5% are not mentioned here.

The alkane hydroxylases are a group of electron carrier dependent monooxygenase systems with the ability to oxidize n-alkanes to their corresponding alkan-1-ol (Li et al., 2008b). This group of enzymes are categorised into different subgroups based on the coenzymes used for hydroxylation of alkanes and the carbon length of substrates (Table 3.6.2). The copper membrane-associated monooxygenases (CuMOs) are a diverse superfamily of multi-subunit enzymes produced by several groups of bacteria and are able to oxidize a range of molecules, including hydrocarbons. Depending on their substrate target these enzymes are called propane monooxygenase (in propane-oxidizing bacteria) (Redmond et al., 2010), methane monooxygenase (in methane oxidizing bacteria) (Luesken et al., 2011) ammonia monooxygenase (AMO; in ammonia-oxidizing bacteria) (Taylor et al., 2013), toluene monooxygenase (in toluene-oxidizing bacteria) (Tinberg et al., 2011) and etc. AMOs are a group of monooxygenase systems with the ability to hydroxylate hydrocarbons through tricarboxylate transport system. They are homologous to many other monooxygenases such as the hydrocarbon monooxygenases (Coleman et al., 2011a, Coleman et al., 2011b), butane monooxygenases (BMOs) (Sayavedra-Soto et al., 2011), and archaeal AMOs (Walker et al., 2010, De La Torre et al., 2008). Approximately 0.002% of the metagenome were recognised as one of these alkane hydroxylase-encoding genes (Table 3.6.2), which belonged to a variety of phylogenetic groups including Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, Deltaproteobacteria, Actinobacteria and several fungi. Analysis on the gene sequences of these genes in NCBI using the Geneious software package showed several novel alkane hydroxylase genes in this crude oil sample. However, since these genes have not been isolated in laboratory conditions, there is no information about their products enzymatic activity and therefore require further studies.

Furthermore, based on this metagenomic study, approximately 1.0% of the whole metabolic gene capacity of microorganisms in crude oil was involved in the degradation of different types of aromatic hydrocarbons (Table 3.6.2). The whole group of metabolic pathways involved in the degradation of aromatic hydrocarbons is indicated in Figure 3.6.5. This analysis showed that 12% of the functional genes involved in degradation of aromatic hydrocarbons are involved in aerobic metabolic pathways. The majority (53%) of the genes responsible for degradation of aromatic compounds in this sample were involved in peripheral pathways for catabolism of different types of aromatic compounds, including benzoate, biphenyl, naphthalene, anthracene, phenol, phenylpropanoid, quinate, salicylate ester, toluene, n-

Phenylalkanoic acid, p-Hydroxybenzoate, p-hydroxybenzoate, cresol, carbazol, p-cymene, gentisate, aromatic-amines and chloroaromatics (Table 3.6.2). Furthermore, approximately 19% of the aromatic degrading genes in this sample were involved in the metabolism of central aromatic intermediates, such as 4-Hydroxyphenylacetic acid catabolic pathway, catechol branch of beta-ketoadipate pathway, central meta-cleavage pathway of aromatic compound degradation, homogentisate pathway of aromatic compound degradation, N-heterocyclic aromatic compound degradation, Protocatechuate branch of beta-ketoadipate pathway and salicylate and gentisate catabolism (Table 3.6.2). In addition to the aerobic pathways, 16% of the genes in this crude oil sample belonged to anaerobic shunts for the degradation of aromatic hydrocarbons such as acetophenone carboxylase 1, anaerobic benzoate metabolism, anaerobic toluene and ethylbenzene degradation and hydroxyaromatic decarboxylase family (Table 3.6.2). A variety of microorganisms from different phylogenetic categories, such as methanogenic consortia, iron-reducing bacteria, denitrifying microorganisms, sulphate reducing bacteria (SRB), and anaerobic phototrophic microorganisms are able to degrade some aromatic hydrocarbons through addition of a carbonic group to a carbon atom (Verfürth et al., 2004, Beasley and Nanny, 2012, Tierney and Young, 2010). The essential enzymes for these catabolic pathways, including benzylsuccinate synthase, Succinyl-CoA:(R)-benzylsuccinate CoA-transferase, (R)-benzylsuccinyl-CoA dehydrogenase, Phenylitaconyl-CoA hydratase, Benzoylsuccinyl-CoA thiolase were found in this data. Another key enzyme for the anaerobic degradation of aromatic hydrocarbons, referred to as ethylbenzene dehydrogenase (EBDH), was identified in this data. Through this catabolic pathway microorganisms are able to degrade ethylbenzene through an anoxic hydroxylation reaction in which a periplasmic EBDH hydroxylates the terminal carbon of the side chain of ethylbenzene to produce S-1-phenylethanol (Heider, 2007, Kniemeyer and Heider, 2001b). This product is further metabolized by a NAD-dependent alcohol dehydrogenase, an acetophenone carboxylase (APC) and benzoylacetate-CoA ligase (BAL) to produce benzoyl-CoA and acetyl-CoA units (Kniemeyer and Heider, 2001b). Anaerobic benzoate metabolism proceeds further using several enzymes such as acetyl-CoA acetyltransferase, Benzoyl-CoA reductase, 2-hydroxycyclohexanecarboxyl-CoA dehydrogenase, 2-ketocyclohexanecarboxyl-CoA hydrolase, Cyclohex-1-ene-1-carboxyl-CoA hydratase, Glutaryl-CoA dehydrogenase and 3-hydroxybutyryl-CoA dehydrogenase, all of which were found in the crude oil metagenome data.

Sulfur-containing hydrocarbons are very common compounds in crude oils which reduce the quality of this product due to the requirement for a considerable amount of further processing, such as enzymatic treatment, hydrotreatment and cracking, to produce more valuable hydrocarbon products (Milam et al., 2014). Microorganisms, however, are able to remove sulfur from these hydrocarbons using several types of enzymes such as alkanesulfonate monooxygenases and dibenzothiophene (DBT) monooxygenases (Wang et al., 2013b, Dayal et al., 2013). Alkanesulfonate monooxygenase converts sulfonated alkanes to alkane and sulfonate. DBT is a polyaromatic sulfur heterocyclic compound that is desulfurized using three enzymes encoded by three ORFs, designated as dszA (Dibenzothiophene sulfone monooxygenase), dszB (2,2-hydroxybiphenyl benzensulfunate desulfinase) and dszC (dibenzothiophene monooxygenase). A remarkable number of these genes (0.005% and 0.49% respectively for alkanosulfate degrading pathway and DBT desulfurizing pathway) and their corresponding transporter genes were found in this metagenomic data, belonging to several microbial strains (Table 3.6.3).

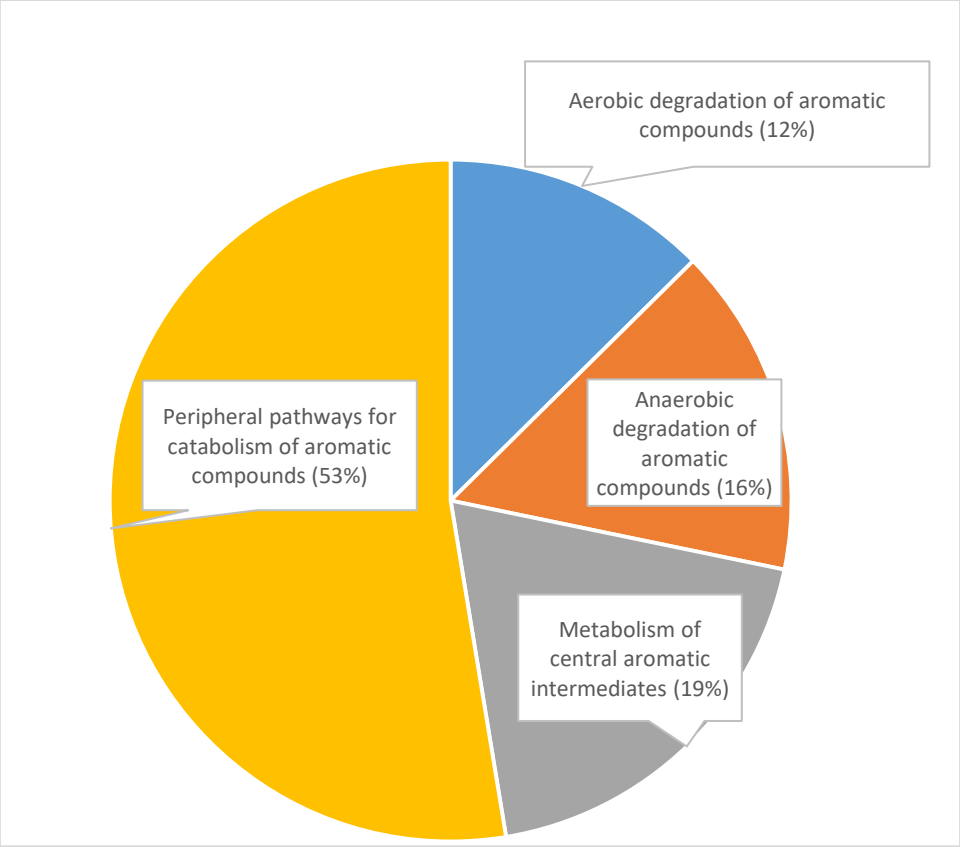


Figure 3.6.5. Ratio of the genes responsible for aromatic degradation in the crude oil metagenome

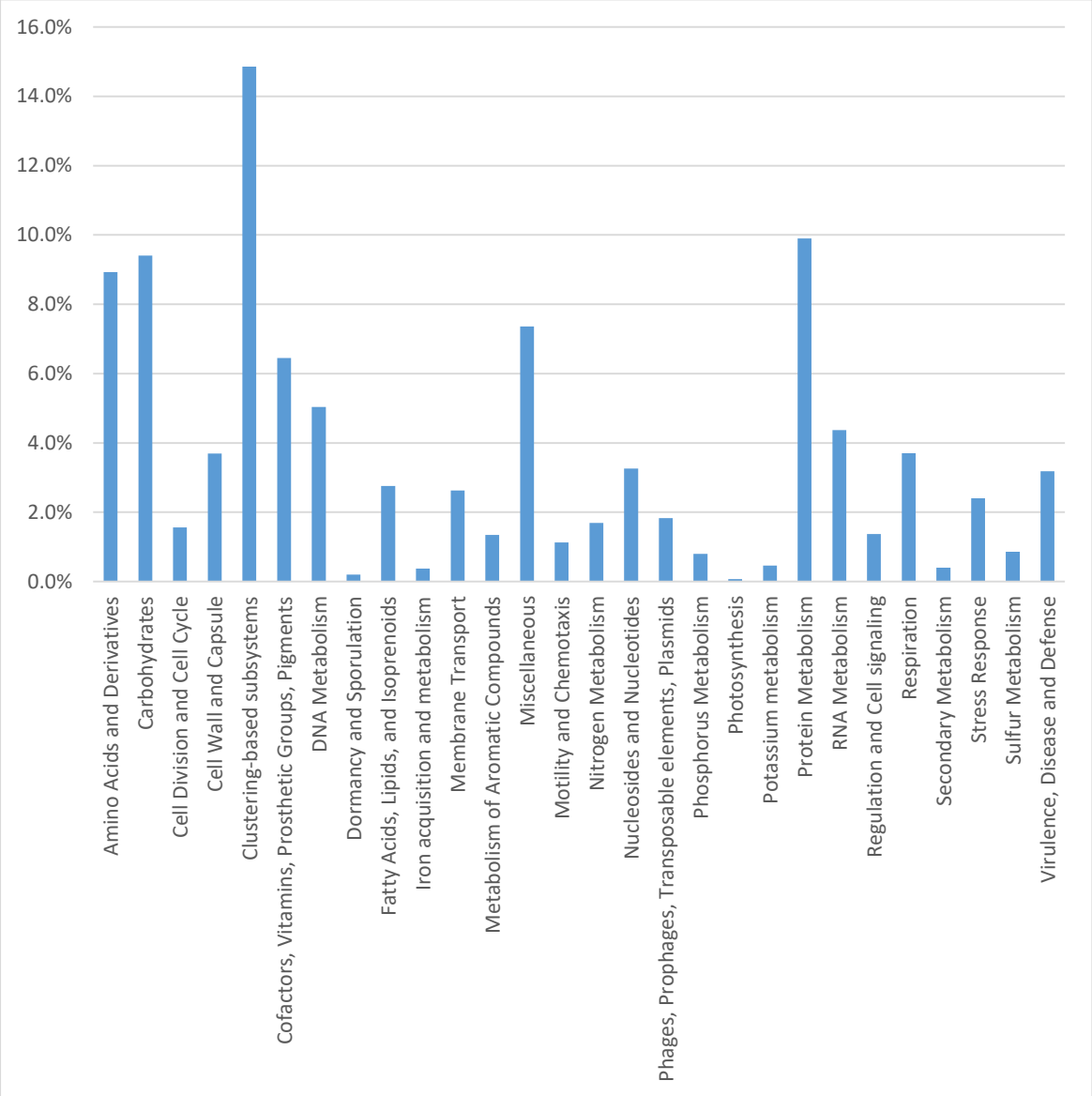


Figure 3.6.4. Ratios of the DNA reads involved in microbial metabolism in crude oil based on annotation of metagenomic data in M5NR source.

Enzymes		Orders	Ratios
Aliphatics	alkane hydroxylase	*Acinetobacter* sp. (*Acinetobacter* sp. ATCC 27244, *A. haemolyticus* ATCC 19194, *A. venetianus*), *Marinobacter* sp. P1-14D, *Pseudomonas frederiksbergensis*, uncultured bacterium	0.0005%
	Propane monooxygenase alpha and beta subunits	*Mycobacterium chubuense* NBB4	0.0003%
	Bacterial Flavin-binding monooxygenase (AlmA; for C20–C36 alkanes)	*Alcanivorax jadensis*, *Aeromonas salmonicida* subsp. salmonicida A449, *Acinetobacter* sp. (*A. junii* SH205, *A. baumannii* ATCC 17978, *Acinetobacter* sp. DSM 17874, *Acinetobacter* sp. RAG-1), *Parvibaculum lavamentivorans* DS-1, *Burkholderia multivorans* ATCC 17616, *Dechloromonas aromatica* RCB, *Ralstonia eutropha* JMP134, *Sorangium cellulosum* So ce 56, *Corynebacterium efficiens* YS-314, *Frankia* sp. EAN1pec, *Marinobacter algicola* DG893, *Mycobacterium* sp. *(M. avium* paratuberculosis K-10, *M. vanbaalenii* PYR-1, *M. marinum* M), *Streptomyces avermitilis* MA-4680, *Salinispora tropica* CNB-440, Stackebrandtia nassauensis DSM 44728, *Nakamurella multipartita* strain ATCC 700099, *Salinispora tropica* CNB-440), *Neurospora crassa*, *Saccharomonospora viridis* strain ATCC 15386, *Magnaporthe oryzae* 70-15, *Rhodococcus jostii* RHA1	0.001%
	Bacterial CYP153 Soluble cytochrome P450 (for C5–C12 alkanes)	*Ajellomyces dermatitidis SLH14081*, *Parvibaculum sp. EPR92*, *Alcanivorax dieselolei Mycobacterium marinum* M, *Dietzia cinnamea* P4	0.0002%
	Methane monooxygenase/ammonia monooxygenase, subunits A, B and C	*Methylosinus trichosporium* OB3b, , *Methylococcus* capsiulatus str. Bath, *Methylocystis* sp. M, *Methylocystis* sp. ATCC 49242, *Methylobacter tundripaludum* SV96, *Nitrosococcus halophilus* Nc4, *Nitrosomonas* sp. AL212, uncultured methanotrophic bacterium, *Citreicella* sp. SE45	0.0005%
	alpha-ketoglutarate-dependent dioxygenase AlkB	*Achromobacter xylosoxidans* A8, *Pectobacterium carotovorum*, *Pantoea ananatis* LMG 20103, *Ralstonia solanacearum* UW551	0.00008%
Aromatics	Aerobic pathways	Aromatic Amine Catabolism, Benzoate transport and degradation cluster, Gentisarate degradation, Phenylacetyl-CoA catabolic pathway (core), Toluene 4-monooxygenase (T4MO), carbazol degradation cluster, p-cymene degradation	0.1%
	Anaerobic pathways	Acetophenone carboxylase 1, Benzoate metabolism, toluene and ethylbenzene degradation, Hydroxyaromatic decarboxylase family	0.2%
	Central aromatic intermediates	4-Hydroxyphenylacetic acid catabolic pathway, Catechol branch of beta-ketoadipate pathway, Central meta-cleavage pathway of aromatic compound degradation, Homogentisate pathway of aromatic compound degradation, N-heterocyclic aromatic compound degradation, Protocatechuate branch of beta-ketoadipate pathway, Salicylate and gentisate catabolism	0.2%
	Peripheral pathways	Benzoate catabolism, Biphenyl Degradation, Chloroaromatic degradation pathway, Naphtalene and antracene degradation, Phenol hydroxylase, Phenylpropanoid compound degradation, Quinate degradation, Salicylate ester degradation, Toluene degradation, n-Phenylalkanoic acid degradation, p-Hydroxybenzoate degradation.	0.45%

Table 3.6.2. List of the genes and microorganisms involved in degradation of aliphatic and aromatic hydrocarbons identified in the metagenomic data analysis of crude oil.

Pathway	Proteins	List of Microorganisms	Ratio
Alkanosulfate degrading pathway	Alkanesulfonate monooxygenase, FMNH(2)-dependent	*Acidovorax avenae* citrulli AAC00-1, *Acinetobacter* sp. (A. *baumannii* 6013113, *A.* sp. ADP1), *Azorhizobium sp. (A. caulinodans* ORS 571, A. *vinelandii* DJ), *B. subtilis* 168, *Bradyrhizobium* sp. BTAi1, *Burkholderia* sp. AUO158, *Burkholderia* sp. 383, *B. ambifaria* AMMD, *B. cenocepacia* AU 1054, *B multivorans* ATCC 17616, *B. xenovorans* LB400, *B. cenocepacia* J2315, *B. xenovorans* LB400), *Cupriavidus metallidurans* CH34, *Delftia acidovorans* SPH-1, *Frankia* sp. EAN1pec, *Granulibacter bethesdensis* CGDNIH1, *Janthinobacterium Marseille*, *Mesorhizobium loti* MAFF303099, *Mycobacterium* sp. (strain JLS), *Nocardia farcinica* IFM 10152, *Nostoc punctiforme* PCC 73102, *Pseudomonas* sp. (*P. entomophila* L48, *p. fluorescens* Pf-5, *P. fluorescens* PfO-1, *P. syringae* pv. tomato DC3000, *P. syringae* pv. syringae B728a, *P. mendocina* ymp), *Ralstonia* sp. (*R. eutropha* JMP134, *R. solanacearum* GMI1000, *R. metallidurans* ATCC 43123), *Rhodoferax ferrireducens* ATCC BAA-621, *Rhodopseudomonas palustris* CGA009, *Rhodococcus jostii* RHA1, *Serratia marcescens* Db11	0.001%
	Alkanesulfonates transport system permease protein	*Acidovorax avenae* citrulli AAC00-1, *Acinetobacter* sp. ADP1, *Acinetobacter baumannii* AB059, *Agrobacterium tumefaciens* C58, *Anaeromyxobacter* sp. Fw109-5, *Azorhizobium caulinodans* ORS 571, *Azotobacter vinelandii* DJ, *B. cereus* BDRD-Cer4, *Bacillus clausii* (strain KSM-K16), *Bradyrhizobium* sp. (strain ORS278), *Bradyrhizobium* sp. BTAi1, *Burkholderia* dolosa AUO158, *Burkholderia multivorans* ATCC 17616, *Burkholderia phymatum* STM815, *Burkholderia xenovorans* LB400, *B. xenovorans* LB400, *Cupriavidus metallidurans* CH34, *Delftia acidovorans* SPH-1, *Desulfotomaculum reducens* MI-1, *Granulobacter bethesdensis* CGDNIH1, *Janthinobacterium* sp. *Marseille*, *Mesorhizobium* loti MAFF303099, *Methylibium petroleiphilum* PM1, *Methylobacterium extorquens* PA1, *Methylobacterium radiotolerans* JCM 2831, *Moorella thermoacetica* (strain ATCC 39073), *P. aeruginosa* (39016, PA7, PACS2), *Pseudomonas entomophila* L48, *P. fluorescens* (Pf-5, PfO-1 and SBW25), *Rhodopseudomonas palustris* HaA2, Pseudomonas mendocina ymp, *P. putida* (GB-1 and W619), *Ralstonia eutropha* JMP134, *Ralstonia solanacearum* GMI1000, *Rhodoferax ferrireducens* T118, *Rhodopseudomonas palustris* HaA2, *Rhodopseudomonas palustris* CGA009, *Serratia proteamaculans* 568, *Sorangium cellulosum* (strain So ce56)), *Thermanaerovibrio* acidaminovorans DSM 6589, *Xylanimonas cellulosilytica* DSM 15894, *Xanthomonas axonopodis* pv. citri str. 306	0.004%
Dibenzothiophene desulfurase	(dszC	*Acidovorax delafieldii*, *Agrobacterium tumefaciens*, *Rhodococcus* sp. XP, *Rhodococcus erythropolis*, Brevibacillus *brevis*, *Gordonia amicalis*, *Gordonia alkanivorans*, *Gordonia* sp. CYKS2, *Mycobacterium avium* 104, *Mycobacterium goodie*, *Starkeya novella* DSM 506	0.17%
	dszA	*Gordonia* sp. (WQ-01A and CYKS2), , *Gordonia alkanivorans*, *Acidovorax delafieldii*, *Brevibacillus brevis*, *Rhodococcus erythropolis*, *Rhodococcus* sp. SDUZAWQ	0.22%
	dszB	*Mycobacterium avium* 104, *Gordonia amicalis*, *Gordonia alkanivorans*, *Gordonia* sp. CYKS2, *Gordonia* sp. F5.25.8, *Rhodococcus erythropolis*, *Brevibacillus brevis*, *Rhodococcus* sp. XP, *Rhodococcus* sp. DS-3, *Acidovorax delafieldii*	0.1%

Table 3.6.3. List of the genes and microorganisms involved in desulfurization of hydrocarbons identified in crude oil metagenomic data analysis; dszC: Dibenzothiophene (DBT) monooxygenase; dszA: DBT-sulfone monooxygenase); dszB: dibenzothiophene-5,5-dioxide monooxygenase.

3.6.5.5 Additional genes involved in microbial survival in crude oil

Several microorganisms express a variety of surface-active compounds made of lipopolysaccharides, lipoproteins and polysaccharides with different functionality including improvement of the bioavailability of water-insoluble molecules, the adherence of the cells to hydrophobic surfaces and probable antimicrobial activity (Soberón-Chávez and Maier, 2011). At least three groups of genes responsible for production of biosurfactants were identified in the crude oil microbiome (mono/di-rhamnolipids, trehalose containing biosurfactants and emulsan). The mono/di-rhamnolipids are a group of exopolysaccharides with the ability to function as surfactants for several types of bacteria (Müller et al., 2012). These surfactants are produced by the action of rhamnosyltransferase 1 and rhamnosyltransferase 1, respectively (Müller et al., 2012), which both were found in the crude oil metagenome data. Furthermore, almost 0.1% of the total protein encoding genes in the crude oil metagenome belonged to the enzymes responsible for the production of trehalose. The disaccharide trehalose is a residue on a glycolipid surfactant which is substituted by mycolic acids at C-6 and C-6' positions (Lang et al., 2000). In addition, presence of UDP-N-acetylglucosamine 2-epimerase, Mannosyltransferase C, Glycosyltransferase, Acetyltransferase and Perosamine synthetase (Nakar and Gutnick, 2001) in this metagenomic data may be used as an indicator of production of emulsans, a group of biosurfactants composed of a repeating N-acetyl-D-galactosamine, N-acetylgalactosamine uronic acid and diamino deoxyhexosamine unit esterified d to fatty acids (Nakar and Gutnick, 2001).

The shortage of sulfur, nitrogen, phosphorus and other elements in crude oil is considered as one of the most important challenges for microbial growth in these habitats (Hasinger et al., 2012a, Megharaj et al., 2011). Based on the metagenomic study, microorganisms in crude oil spared a significant portion of their functional genomic capacity to the genes responsible for acquisition of sulfur, nitrogen, phosphorus, iron and potassium (approximately 0.7%, 0.15%, 1.1%, 0.4% and 0.4%, respectively) (Table 3.6.4). On the other hand, several studies have confirmed the consistent presence of different heavy metals, such as arsenic, mercury, manganese, chromium, lead, molybdenum, cobalt, copper, nickel, vanadium and cadmium in crude oils (Erickson et al., 1954, Osuji and Onojake, 2004, Gondal et al., 2006). Therefore, resistance to these heavy metals is a critical factor for microbial survival in these environments. In accordance with these hypothesis, our data from metagenomic analysis showed that several genes detected in in crude oil were involved in resistance to many heavy metals such as arsenic, cadmium, cobalt, zinc, copper and mercury (0.2%, 0.06%, 0.2%, 0.3% and 0.01%,

respectively) (Table 3.6.4). These genes are responsible for the excretion of heavy metals (such as arsenic efflux pump, mercuric transport protein, cusCFBA proton-cation antiporter complexes, *copA* ATPase dependent transporter, Cadmium-transporting ATPase, Co/Zn/Cd efflux system membrane fusion protein) or their conversion to compounds or valance states with lower toxicity, such as arsenate reductase, mercuric reductase, organomercurial lyase, multicopper oxidase).

3.6.6 Conclusion

Crude oil as used in this study seems to offer a suitable environment for growth of a diverse community of microorganisms belonging to several different groups of bacteria, including Proteobacteria, Deferribacteres, Bacteroidetes, Firmicutes, Cyanobacteria, Actinobacteria, Euryarchaeota and many others. These microorganisms possessed several enzymes that could be used to degrade different aliphatic and aromatic hydrocarbons. Furthermore, these microorganisms possessed a variety of mechanisms that could be used to cope with the extreme conditions of this environment, including the C:N:P imbalance between carbon sources and other nutrients , the hydrophobicity of hydrocarbons and toxicity of heavy metals. Overall, such information can be very useful for understanding the physiological niche of various microorganisms in crude oil. In terms of environmental biotechnology, the biodegradation of hydrocarbons and clearance of crude oil spills have been big environmental concerns. Metagenomic data is very helpful in allowing the development of optimal conditions for the growth and activity of hydrocarbon degrading microorganisms in such environments. Furthermore, such studies provide much useful information on those microorganisms with the ability to remove many elements from oil thus helping in the refining of this product. In addition, a deep study on the biosurfactant producing microorganisms in crude oil may improve the production of biosurfactants applicable in many industries.

Ability	Elements	Mechanism	Ratios
Metal resistance genes	Arsenic	Arsenate reductase (EC 1.20.4.1), Arsenic efflux pump protein, Arsenic resistance protein ArsH, Arsenical pump-driving ATPase (EC 3.6.3.16), Arsenical resistance operon repressor, Arsenical resistance operon trans-acting repressor ArsD, Arsenical-resistance protein ACR3	0.17%
	Cadmium	Cadmium efflux system accessory protein and Cadmium resistance protein	0.06%
	Cobalt-zinc-cadmium	Cadmium-transporting ATPase (EC 3.6.3.3), Cation efflux system protein CusA, Cd(II)/Pb(II)-responsive transcriptional regulator, Cobalt-zinc-cadmium resistance proteins (CzcA, CzcC and CzcD), Cobalt/zinc/cadmium efflux RND transporter, membrane fusion protein, CzcB family, Copper sensory histidine kinase CusS, Copper-sensing two-component system response, regulator CusR, DNA-binding heavy metal response regulator, Heavy metal RND regulators (efflux outer membrane protein, transcriptional regulator HmrR), sensor histidine kinase, Hypothetical protein, Transcriptional regulator, MerR family	0.2%
	Copper	Blue copper oxidase CuO precursor, CopG protein, Copper chaperone, Copper resistance proteins (B, C and D), Copper sensory histidine kinase (CusS) and regulator (CusR), Copper-translocating P-type ATPase, Cu(I)-responsive transcriptional regulator, Cytochrome c heme lyase (subunits CcmF and CcmH), Heavy metal-(Cd/Co/Hg/Pb/Zn)-translocating P-type, ATPase: Heavy metal translocating P-type ATPase, Multicopper oxidase, Multidrug resistance transporter, Bcr/CflA family, Sensor protein copS, Transcriptional activator protein CopR, Copper homeostasis protein CutE, Magnesium and cobalt efflux protein CorC, Secreted protein, suppressor for copper-sensitivity ScsC,	0.33%
	Mercury	Mercuric ion reductase (EC 1.16.1.1), PF00070 family, FAD-dependent NAD(P)-disulphide, oxidoreductase, Mercuric ion reductase (EC 1.16.1.1), Mercuric resistance operon coregulator, Mercuric resistance operon regulatory protein, Mercuric transport protein, MerC, Mercuric transport protein, MerE, Mercuric transport protein, MerT, Organomercurial lyase (EC 4.99.1.2), Periplasmic mercury(+2) binding protein	0.01%
	Chromium	Chromate resistance protein ChrB, Chromate transport protein ChrA, Superoxide dismutase ChrC, Superoxide dismutase SodM-like protein ChrF	0.03%
	Zinc	Response regulator of zinc sigma-54-dependent two-component system, Zinc resistance-associated protein	0.22%
Genes involved in element shortage	Sulfur	Galactosylceramide and sulfatide metabolism (alpha-galactosidase, alpha-galactosidase precursor, arylsulfatase, beta-galactosidase and neuraminidase), sulfate reduction-associated complexes (glutamate synthase, sulfite reductase, HmeA, HmeF, HmeD, HmeC and tRNA 2-thiouridine synthesizing protein E), sulfur oxidation (Blr3520 protein homolog, Lipocalin-related protein and Bos/Can/Equ allergen, Sox proteins, sulfur oxidation molybdopterin C protein), thioredoxin-disulfide reductase (Alkyl hydroperoxide reductase protein), inorganic sulfur assimilation and organic sulfur assimilation (Alkanesulfonates utilization, DMSP breakdown, L-Cystine uptake and metabolism, taurine utilization and utilization of glutathione as a sulphur source	0.65%
	Nitrogen	Allantoin utilization, ammonia assimilation, nitrogen fixation, cynate hydrolysis, denitrification, dissimilatory nitrite reductase, nitrate and nitrite ammonification and nitric oxide synthase	0.15%
	Phosphorus	Alkylphosphonate utilization, high affinity phosphate transport, phosphate-binding DING proteins, alkaline phosphatase, exopolyphosphatase, inorganic pyrophosphatase, low-affinity inorganic phosphate transporter, NAD(P) transhydrogenase, PhoQ, phosphatase, phosphate ABC transporter, phosphate transport system permease proteins (PstA and PstC), polyphosphate kinase, pyrophosphate-energized proton pump, sodium-dependent phosphate transporter, soluble pyridine nucleotide transhydrogenase, 2-aminoethylphosphonate:pyruvate	1.1%

		aminotransferase and 2-aminoethylphosphonate ABC transporter permease protein II)	
	Iron	Production of siderophores and their uptake, ABC-type iron transport system, heme and hemin uptake and utilization systems, iron-binding protein, permease proteins, nonheme iron-containing ferritin, Iron(III) dicitrate transport system Fec	0.43%
	potassium	Glutathione-regulated potassium-efflux system, includingATP-binding protein, ancillary proteins (KefF, KefG, KefB, KefA and potassium efflux system YjeP) and Potassium homeostasis (consisting of several types of potassium pump systems)	0.37%

Table 4.6.4. Ratios of the genes involved in metal resistance and shortage of elements identified in crude oil metagenomic data.

3.7 Identification of a new operon involved in desulfurization of Dibenzothiophenes using a metagenomic study and cloning and functional analysis of the genes

Published in: Enzyme and Microbial Technology; 87: 24-28.

3.7.1 Abstract

The presence of sulphur-substituted hydrocarbons in fossil fuels are one of main reasons for the release of sulfur oxides into the environment. Dibenzothiophenes (DBT) are organic sulfur-containing molecules in crude oil, which have the potential for biological oxidation, with the sulphur being removed through an enzymatic cleavage of the C-S bonds. Therefore, finding new strains that can desulfurize this compound has recently become a point of interest. In this study, three new genes involved in the bacterial desulfurization of Dibenzothiophene, which were sequenced in the course of a metagenomic study, were isolated by PCR amplification in the laboratory. The activities of these genes were then analysed following insertion into an expression vector and cloning in *E.coli* DH5α cells. Based on the results, all three genes were actively expressed and their products could act on their corresponding substrates.

3.7.2 Introduction

Fossil fuels, including crude oil and coal, are both the main source of energy and the main feed stock for the petrochemical industries, and therefore their usage in industries constantly increases (Abbasian et al., 2015a). However, these fossil fuels contain high levels of nitrogen and sulphur-substituted hydrocarbons (Masnadi et al., 2014). The level of sulfur content in crude oils varies between 0.03% to 7.89% (w/w) (Le Borgneb, 2004). The combustion of these types of hydrocarbons leads to the production and release of nitrogen/sulfur oxides, which are considered as the main source for the formation of acid deposition (Khan et al., 2009). Although there are many methods, such as post-combustion desulfurization techniques, for the removal of these elements from these fossil fuels, they are expensive and inconvenient (Soleimani et al., 2007). Furthermore, while inorganic sulfur compounds, including elemental sulfur, sulphide, sulphite, sulfate and thiosulfate, are easily removed from crude oil and coal using physical and chemical treatments (Meyer, 2013), removal of most of organic sulfur containing compounds, such as dibenzothiophene (DBT), benzothiophene and their derivatives is not easily achieved (Soleimani et al., 2007, Lu et al., 2007). These concerns and increased

considerations regarding the acid deposition issues, has increased the interest in the use of microbial sulfur transformation reactions.

Dibenzothiophene and its derivatives, are considered as the prototype for the organic sulfur-containing molecules in crude oil, and much effort has already been put into the development of a biological oxidative approach to the removal of sulfur from these compounds using the enzymatic cleavage of the C-S bonds (Lu et al., 2007). The majority of DBT desulfurizing microorganisms and their DBT desulfurizing operons have been found using traditional culture and molecular based approaches (Bhatia and Sharma, 2012, Seo, 2012, BUZANELLO et al., 2014, Bahuguna et al., 2011, Dai et al., 2014, Bordoloi et al., 2014, Shi et al., 2014, Chauhan et al., 2014). However, these systems are not able to identify the majority of these microorganisms in natural sources (Mocali and Benedetti, 2010). Metagenomic approaches, however, are theoretically able to show the whole genome contents of the microbial community present in an environment (Abbasian et al., 2015b). In a metagenomic study on crude oil, we obtained a list of microorganisms and the genes involved in DBT desulfurization (unpublished data). Although many of these genes have already been identified and can be found in public data bases, a few were quite different. In this study, we focussed on one of these operons in order to obtain its full sequence and to study the desulfurizing activity of the enzymes produced.

3.7.3 Material and Methods

3.7.3.1 Gene sequence and Primer design

The original gene fragment obtained from the metagenomic study was expanded using Geneious software and reiterative mapping of short reads from the whole metagenomic data in order to obtain the full length operon sequence. This operon sequence was analysed using Geneious software to find suitable primers for amplification through PCR reactions.

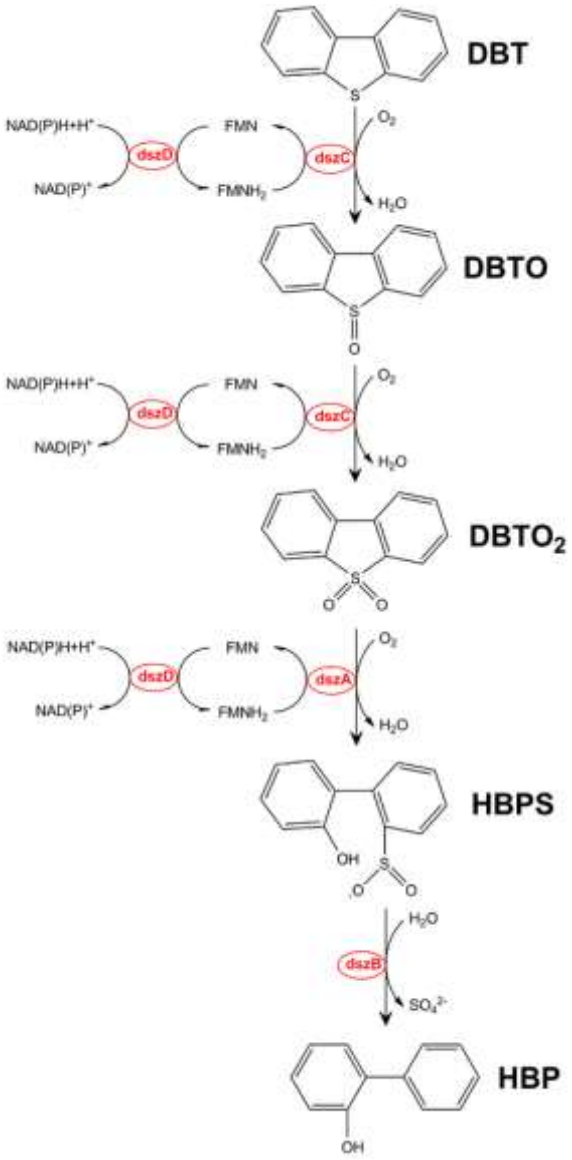


Figure 3.7.1. The Dibenzothiophene desulfurizing pathway in *Gordonia* sp; dszA: Dibenzothiophene sulfone ($DBTO_2$) monooxygenase, dszB: 2-(2-hydroxyphenyl)-benzene sulfonate (HBPS) desulfinase, dszC: Dibenzothiophene monooxdase and dszD: NAD(P)H:FMN oxidoreductase, DBT: Dibenzothiophene, DBTO: Dibenzothiophene-sulfoxide, $DBTO_2$: Dibenzothiophene sulfone, HBPS: 2-(2-hydroxyphenyl)-benzene sulfonate, HBP: 2-Hydroxybiphenyl.

3.7.3.2 Gene cloning and enzyme assay

DNA extractions was conducted on 100 ml crude oil using the power water DNA kit (MO BIO), according to the manufacturer's standard instructions. This volume of crude oil was

filtered using a disposable filter (0.22 µm) to isolate and concentrate the microbial cells. The microorganisms on the filters were dissolved into 1 ml of solution PW1 (provided by the manufacture), and further DNA extraction processes were performed according to their instructions. PCR amplification reactions for each gene were performed separately using 1µM of the primers designed by our team. The PCR products were purified by electrophoresis on agarose gels and the DNA from the purified bands was ligated with the expression vector pGEM-T Easy (Promega Madison, USA) vector according to the manufacture's instruction. The ligated plasmids were transformed into DH5α cells (provided from Bioline) using standard techniques and plated on plates allowing blue/white selection of recombinant plasmids. The white colonies obtained from growth of bacteria on Luria–Bertani (LB) medium supplemented with ampicillin ($100\ mg^{-1}$) and X-gal were isolated and the inserts sequenced with the universal M13 primers (M13F: CGCCAGGGTTTTCCCAGTCACGAC and M13R: TCACACAGGAAACAGCTATGAC) (Asari et al., 2013) (using the SA Pathology sequencing facility at Flinders University, Adelaide, Australia).

E. coli DH5α cells were employed as the host for all three genes. The cells were grown in LB broth supplemented with ampicillin (100 mg/ml) under 37 ºC, 200 rpm for an overnight. When the cell absorbance reached 0.5, gene expression was induced by addition of IPTG for 2 hours. These cell suspensions were centrifuged and the pellets were washed twice with 0.1M potassium phosphate buffer (PBS) (pH7.0). The pellets were then re--suspended into the same PBS buffer. These cell suspensions were lysed using sonication, and the supernatants were separated from cell debris using centrifugation (10,000×g, 4 °C) for 5 minutes. Following addition of 9 mM NADH and 0.01mM FMN to the supernatants, the concentration of cell-free lysates was adjusted to 2.0 mg/ml. All three substrates, dibenzothiophene (DBT), dibenzothiophene sulfone ($DBTO_2$) and 2-(2-hydroxyphenyl)-benzene sulfonate (HBPS), were separately dissolved as dimethylformamide solutions to a volume of 100 mM, and were mixed with an adequate of cell-free lysate mixtures to a final concentration of 1mM. The enzyme activity of the DszA and DszC cloned genes were assayed using absorbance of NADH at 340 nm for 2 hours (using UV.360-Shimadzu Spectrophotometry) in separate suspensions containing DszA and DszC recombinant *E. coli* DH5α and the corresponding substrates, and their comparison with this value in the suspensions containing the non-recombinant *E. coli* DH5α and the corresponding substrate for each enzyme. The enzyme activity of DszB was

determined based on pH changes in the suspension containing DszB recombinant *E. coli* DH5α and substrate and comparison with this value in the non-recombinant *E. coli* DH5α.

3.7.3.3 Accession number

The nucleotide sequence data for these three desulfurization genes, were submitted to the GenBank databases under the accession numbers KT630579, KT630580 and KT630581 for genes ADRO1, ADRO2 and ADRO3 respectively.

3.7.3.4 Results and discussion

The dibenzothiophene desulfurizing bacteria harbour an operon containing three genes (DszA, DszB and DszC), which perform the desulfurization process in four steps though the 4S desulfurization pathway (Raheb, 2011, Shavandi et al., 2013) (Fig. 3.7.1). The first two reactions, performed by Dibenzothiophene monooxidase (DszC) convert DBT first to dibenzothiophene-sulfoxide and then to dibenzothiophene sulfone ($DBTO_2$) (Abin-Fuentes et al., 2013, Martinez et al., 2015). This intermediate product is later catalysed by $DBTO_2$ monooxygenase (DszA) to produce 2-(2-hydroxyphenyl)-benzene sulfonate (HBPS) (Abin-Fuentes et al., 2013, Martinez et al., 2015). The product of DszB gene (HBPS-desulfinase) releases sulfur from the compound and leaves 2-Hydroxybiphenyl (HBP). The first three steps are O_2-dependent oxidative reactions and require an electron and hydrogen transportation system, composed of $FMNH_2$ and NAD(P)H as a reductant (Abin-Fuentes et al., 2013, Martinez et al., 2015). Several gene fragments responsible for DBT desulfurization were found in a metagenomic study on crude oil and crude oil field soils (unpublished data). Although many of these fragments showed high similarity with the genes deposited in gene databases, several new sequences were also discovered in these datasets. Gene walking expansion of these fragments using Geneious software to reiteratively map short reads from the metagenomic data produced whole gene sequences from these fragments. Several of these gene sequences showed quite low degrees of similarities to the genes present in public databases. Further expansion of one of the gene sequences in this study produced a whole operon potentially encoding the genes responsible for DBT desulfurization. This operon carried three DBT desulfurizing genes, referred to as *dszA*, *dszB* and *dszC*, which showed low similarities to genes deposited in databases; *dszA* with 78% similarity to dibenzothiophene desulfurization enzyme A of *Gordonia* sp. RIPI, *dszB* with 73% similarity to 2'-hydroxybiphenyl-2-sulfinate desulfinase of *Gordonia* sp. SYKS2 and *dszC* with 77% similarity to dibenzothiophene desulfurizing enzyme

C of *Gordonia* sp. RIPI (Akhtar et al., 2015). These new DBT desulfurizing genes contained 1167 bp (*dszA*), 1116 bp (*dszB*) and 1242 bp (*dszC*) including initiation and stop codons and were nominated as ADRO1, ADRO2 and ADRO3 in the gene bank. The big differences in the similarity of sequences and the nucleotide numbers of these three genes in comparison to other corresponding dsz genes present in gene data banks indicated the existence of new sets of dsz genes belonging to uncultured microorganisms. The presence of potentially novel varieties of the *dsz* genes has been observed by several other researches (Rhee et al., 1998, Bordoloi et al., 2014).

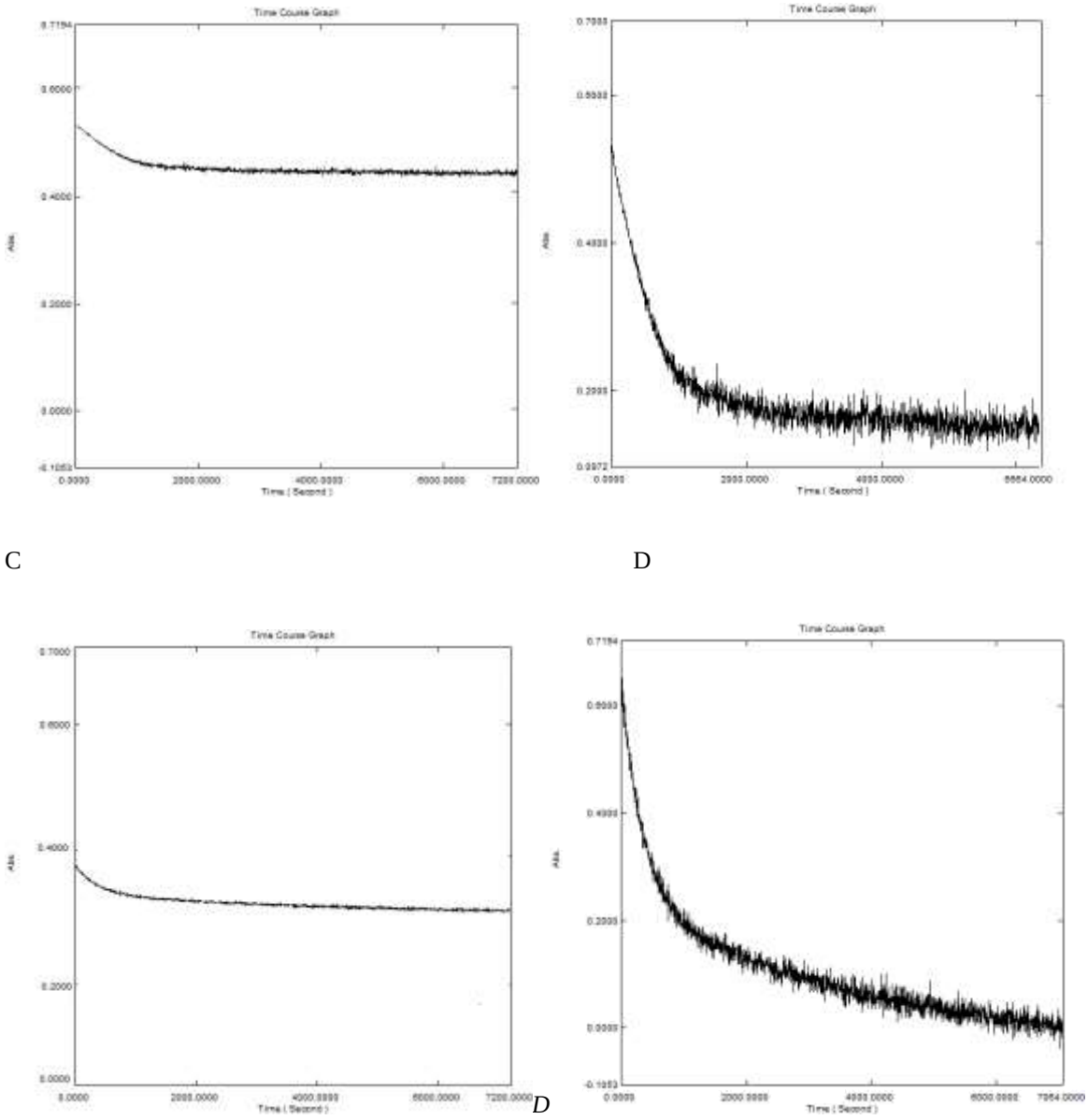
Following construction of three sets of specific primers for the sequences upstream of the initiation codon and downstream of the stop codon of each gene, all three genes were amplified separately by PCR amplification from the DNA obtained from crude oil and crude oil field soil. The presence and amplification of all three genes in the samples were confirmed based on the size of bands on agarose gels. These bands were purified and ligated into the expression vector pGEM-T, and the resulting plasmids transformed into *E. coli* DH5α cells. The presence of genes in these clones were confirmed using the size of plasmids obtained from the transformed bacteria and the sequencing of these plasmids using universal M13 primers. Sequencing of all three genes produced sequences similar to the sequences created originally by software-based assemblage of overlapping reads from the metagenomic data. The predicted molecular weight of the products of these three genes were 96.38 KDa, 91.16 kDa and 102.1 kDa, respectively. There were a 192 bp distances between the stop codon of ADRO1 and initiation codon of ADRO2 (Fig. 3.7.2). Furthermore, a 4 bp overlap was found between the stop codon of second ORF and initiation codon of third ORF in this operon where the initiation codon of ADRO3 started 1 bp before the stop codon of ADRO2 (Fig. 3.7.2). The presence or overlaps or gaps between these genes and their length are variable in different strains. For example, while in *Paenibacillus* sp. A11-2 (Ishii Y et al., 2000) and *Bacillus subtilis*WUS2B (Kirimura K et al., 2004), the initiation codon of *dszB* is located 4 bp before the stop codon of *dszA*, this overlapping region reached 64 in *G. alkanivorans* strain 1B where there is a 64 bp overlap between *dszA* and *dszB* (Alves L et al., 2007).

Figure 3.7.2. The length and direction of genes involved in Dibenzothiophene desulfurizing pathway found in this study, and the gaps and overlaps of these three genes. These three genes (*dszA*, *dszB* and *dszC*), showed 78%, 73% and 77% similarities, respectively, with the corresponding genes in comparison with the submitted genes in NCBI.

The enzyme productivity and activity of these three recombinant *E. coli* DH5α cells were examined using cell-free lysates of each recombinant strain after IPTG induction. The time course of NADH oxidation by each recombinant *E. coli* DH5α strain and the non-recombinant *E. coli* DH5α parent growing in LB media with the corresponding substrates was measured for 2 hours. Based on Beer's Law, the changes in the absorbance of NADH is linearly equal to reduction of the concentration of substrate, and therefore is representative of enzyme activity (Matsubara et al., 2001). Since these genes were expressed under the *lacZ* promoter, their products were produced continuously after IPTG induction (data not shown). The NADH absorbance in the assay with crude Dibenzothiophene monooxdase (dszC) dropped from 0.5 to a number closed to zero after 34 minutes (Fig. 3.7.3) with a rate of 2 µM reduction of substrate per each millimolar enzyme. Furthermore, the NADH absorbance in the assay with crude Dibenzothiophene sulfone (dszA) dropped from 0.6 to a zero after 100 minutes (Fig. 3.7.3), with a rate of 1 µM reduction of substrate per each millimolar enzyme. These reactions continued to make a flat curve after 30 minutes. However, the non-recombinant control cells showed a completely flat curve, which showed that the activity was due to the cloned genes. While the pH value of the suspensions containing DszB recombinant cells and its corresponding substrate changed from 7.4 to 5.8, there were slight changes in this value (7.3)

in the same suspensions containing non-recombinant *E. coli* DH5α strains. This drop in the pH value occurred due to the activity of the products of dszB gene, and therefore, release of sulfuric acid through reaction of HPBS to HBP. As a conclusion, in this study, a new operon responsible for desulfurization genes was identified from environmental samples contaminated with crude oil and were successfully cloned into *E. coli* cells. This operon, consisting of dszA, dszB and dszC, belonged to an uncultured bacterium. These genes were successfully expressed in the laboratory conditions, and therefore, are potentially useful in biodesulfurization of fossil fuels.

A B



C D

D

Figure 3.7.3. The spectrophotometric graphs of enzyme activity based on the levels of NADH oxidation with time. While the levels of NADH oxidation in the dszC and dszA recombinant strains (B and D in the figure, respectively) dropped to a concentration of close to zero, these values did not change significantly in the corresponding control non-recombinant strains (A and C, respectively).

Chapter 4. Summary, conclusions and outlooks

Environmental contamination with crude oil and its products is a serious global environmental issue, and bioremediation using microorganisms is a promising approach to decrease the toxic effects of these compounds. Bioremediation is an effective method which can convert toxic contaminants into biomass or safe or less hazardous compounds. Microorganisms are able to degrade the hydrocarbons in crude oil and its derivatives using enzymatic digestion, and the products are uptaken by the microorganisms or are converted to simplest compounds which may evaporate from the system. Since the presence of microorganisms are compulsory for oil biodegradation, characterization of effective microbial communities in the crude oil and polluted sites and their enzymes involved in hydrocarbon degradation are necessary for development of a microbial strains as pure or mixed cultures for bioremediation. At the same time, these microorganisms and their enzymes can be applied for removal of contaminant elements, such as sulfur and nitrogen, from crude oil, which can be used for improvement of the quality of crude oil.

As first aim of this study, we used metagenomics to investigate the presence of known and probably potential new microorganisms and the genes responsible for degradation of hydrocarbons. While there was differences in the type of dominant microorganisms in different habitats, this study showed that the majority of microorganisms in crude oil and crude oil polluted site belonged to Actinomycetales, Clostridiales, Bacillales, Methanocellales, Flavobacteriales, Pseudomonadales, Xanthomonadales, Burkholderiales, Rhizobiales, Sphingomonadales and Bacterioidales. Furthermore, the majority of these catabolic genes in crude oil and the crude oil well fields were present in Bacteroidetes, Actinobacteria, Firmicutes, Bacteroidetes, Deferribacteres, Planctomycetes and Acidobacteria.

The alkane monooxygenase genes were found in a variety of phylogenetic groups, including Actinomycetales, Burkholderiales, Pseudomonadales, Oceanospirillales, Deinococcales, Rhodobacterales and Sphingobacteriales. Furthermore, the soluble methane monooxygenases (sMM), which are able to hydroxylate hydrocarbons in addition to oxidation of methane, were detected in *Acidiphilium* sp., *Amycolatopsis* sp., *Beijerinckia* sp., *Crenothrix* sp., *Frankia* sp., *Marinobacter* sp., *Mycobacterium smegmatis* str. MC2 155, *Verminephrobacter eiseniae* EF01-2 and many ammonia oxidising bacteria. The gene sequences of many of the alkane monooxygenases found in this study were highly different from the sequences recorded in gene

databases, and could be reported as new genes. However, further studies are required to identify their abilities. In addition, almost big portions (1% and 2.5%) of the genes found in the metagenomic analysis of the crude oil and the soil samples, respectively, were responsible for degradation of aromatic hydrocarbons, including benzene, toluene, carbazol, naphthalene and Ethylbenzene through aerobic and anaerobic metabolism. Another functional group of genes in these metagenomes were responsible for desulfurization of sulfur-containing aliphatic and aromatic hydrocarbons. The alkanesulfonates monooxygenase, which remove sulfur from sulfonated alkanes and dibenzothiophene (DBT) belonged mostly to Actinomycetales, Clostridiales, Burkholderiales, Pseudomonadales and Rhizobiales. Since the DNA sequences of genes give no idea of their expression and enzymatic activities, further studies are required to isolate and express these genes in the laboratory.

In addition to the functional degrading enzymes, these metagenomes contained several sets of genes responsible for adaptation of microorganisms to the harsh environment governed by crude oil. Several genes in these metagenome data were related to production of different types of bio-surfactants (such as mono/di-rhamnolipids, trehalose containing biosurfactants and emulsan). In addition, several genes in these metagenomic data were categorised as the gene groups responsible for resolving the pressure caused by the shortage of sulfur, nitrogen, phosphorus and other elements in crude oil. Furthermore, a number of genes in these data were responsible for microbial resistance to different heavy metals, such as arsenic, mercury, manganese, chromium, lead, molybdenum, cobalt, copper, nickel, vanadium and cadmium found normally in crude oils. These genes belonged mostly to active and inactive cell membrane integrated channels, and the genes responsible for reducing the toxicity of the heavy metals through their oxidation, reduction or conjugation with a (in)organic compound.

This study has focused on different functional genes responsible for degradation of hydrocarbons to find any nobility in these metagenomic data. Several new genes were found in this study which belonged to different classes of hydrocarbon hydroxylating or desulfurizing enzymes. The sequences of many of these genes were submitted in public gene databases. A DNA fragment in this metagenomic data was found as a hydrocarbon desulfurizing enzyme used normally for removal of dibenzothiophene (DBT), the prototype for the organic sulfur-containing molecules in crude oil. Software based-gene walking expansion of these fragments could present whole operon of that fragment, containing *dszA*, *dszB* and *dszC* genes. NCBI

blasting of this genes showed quite low degrees (73-78%) of similarities to the genes present in public databases, and those belonged to an uncultured bacterium. Following design of primer sets for each genes, referred to as ADRO1, ADRO2 and ADRO3 respectively, all these three genes were amplified separately by PCR amplification from the DNA obtained from crude oil and crude oil field soil, and following cloning in *E. coli* DH5α cells, their activity was investigated using expression protocols.

Outlook

For the purpose of future development of this area, the following suggestions can be outlined:

1. Further metagenomic analysis on the new genes found in this study to identify more new hydrocarbon degrading genes.

2. To apply the strains of microorganisms or a combination of those microorganisms found in this study in a field based bioremediation to investigate their ability for removal of crude oil contamination.

3. To apply the enzymes found in this study for improvement of the quality of crude oil in terms of removal of sulfur and nitrogen.

References:

AAGAARD, K., PETROSINO, J., KEITEL, W., WATSON, M., KATANCIK, J., GARCIA, N., PATEL, S., CUTTING, M., MADDEN, T. & HAMILTON, H. 2013. The Human Microbiome Project strategy for comprehensive sampling of the human microbiome and why it matters. *The FASEB Journal,* 27**,** 1012-1022.

ABBAI, N. S., GOVENDER, A., SHAIK, R. & PILLAY, B. 2012. Pyrosequence analysis of unamplified and whole genome amplified DNA from hydrocarbon-contaminated groundwater. *Molecular biotechnology,* 50**,** 39-48.

ABBASIAN, F., LOCKINGTON, R., MALLAVARAPU, M. & NAIDU, R. 2015a. A Comprehensive Review of Aliphatic Hydrocarbon Biodegradation by Bacteria. *Applied Biochemistry and Biotechnology,* 176**,** 670-699.

ABBASIAN, F., LOCKINGTON, R., MALLAVARAPU, M. & NAIDU, R. 2015b. The Integration of Sequencing and Bioinformatics in Metagenomics. *Reviews in Environmental Science and Bio/technology,* 14**,** 357-383.

ABBASIAN, F., LOCKINGTON, R., MALLAVARAPU, M. & NAIDU, R. 2015c. A pyrosequencing-based analysis of microbial diversity governed by ecological conditions in the Winogradsky column. *World Journal of Microbiology and Biotechnology,* 31**,** 1115-1126.

ABBASIAN, F., LOCKINGTON, R., MALLAVARAPU, M. & NAIDU, R. 2015d. A pyrosequencing-based analysis of microbial diversity governed by ecological conditions in the Winogradsky column. *World Journal of Microbiology and Biotechnology,* 31**,** 1115-26.

ABBASIAN, F., LOCKINGTON, R., MEGHARAJ, M. & NAIDU, R. 2016a. The Biodiversity Changes in the Microbial Population of Soils Contaminated with Crude Oil. *Current Microbiology***,** 1-8.

ABBASIAN, F., LOCKINGTON, R., MEGHARAJ, M. & NAIDU, R. 2016b. A Review on the Genetics of Aliphatic and Aromatic Hydrocarbon Degradation. *Applied biochemistry and biotechnology,* 178**,** 224-250.

ABBASIAN, F. & SABERBAGHI, T. 2013a. Metagenomic study of human gastrointestinal tracts in health and diseases. *Gastroenterology and Hepatology Research,* 2**,** 885-896.

ABBASIAN, F. & SABERBAGHI, T. 2013b. Metagenomic Study of Human Gastrointestinal Tracts in Health and Diseases. *Journal of Gastroenterology and Hepatology Research,* 2.

ABIN-FUENTES, A., MOHAMED, M. E.-S., WANG, D. I. & PRATHER, K. L. 2013. Exploring the mechanism of biocatalyst inhibition in microbial desulfurization. *Applied and environmental microbiology,* 79**,** 7807-7817.

ABU LABAN, N., SELESI, D., RATTEI, T., TISCHLER, P. & MECKENSTOCK, R. U. 2010. Identification of enzymes involved in anaerobic benzene degradation by a strictly anaerobic iron-reducing enrichment culture. *Environmental microbiology,* 12**,** 2783-2796.

ACLAND, A., AGARWALA, R., BARRETT, T., BECK, J., BENSON, D. A., BOLLIN, C., BOLTON, E., BRYANT, S. H., CANESE, K. & CHURCH, D. M. 2013. Database resources of the National Center for Biotechnology Information. *Nucleic acids research,* 41**,** D8-D20.

AECKERSBERG, F., BAK, F. & WIDDEL, F. 1991. Anaerobic oxidation of saturated hydrocarbons to CO2 by a new type of sulfate-reducing bacterium. *Archives of Microbiology,* 156**,** 5-14.

AGRAWAL, P., LATHA, S. & MAHADEVAN, A. 1997. Utilization of phenylalanine and phenylacetic acid byPseudomonas solanacearum. *Applied biochemistry and biotechnology,* 61**,** 379-391.

AHMAD, A., MUJEEB, M., KAPOOR, R. & PANDA, B. P. 2013. In situ bioconversion of compactin to pravastatin by Actinomadura species in fermentation broth of Penicillium citrinum. *Chemical Papers,* 67**,** 667-671.

AKBARI, A. & GHOSHAL, S. 2015. Effects of Diurnal Temperature Variation on Microbial Community and Petroleum Hydrocarbon Biodegradation in Contaminated Soils from a Sub-Arctic Site. *Environmental microbiology*.

AKHTAR, N., GHAURI, M. A., ANWAR, M. A. & HEAPHY, S. 2015. Phylogenetic characterization and novelty of organic sulphur metabolizing genes of Rhodococcus spp.(Eu-32). *Biotechnology letters,* 37**,** 837-847.

AL-MAILEM, D., KANSOUR, M. & RADWAN, S. 2015. Moderately thermophilic, hydrocarbonoclastic bacterial communities in Kuwaiti desert soil: enhanced activity via Ca2+ and dipicolinic acid amendment. *Extremophiles***,** 1-11.

ALBUQUERQUE, L., SIMOES, C., NOBRE, M. F., PINO, N. M., BATTISTA, J. R., SILVA, M. T., RAINEY, F. A. & DE COSTA, M. S. 2005. Truepera radiovictrix gen. nov., sp.

nov., a new radiation resistant species and the proposal of Trueperaceae fam. nov. *FEMS microbiology letters,* 247**,** 161-169.

ALIOTO, T., PICARDI, E., GUIGÓ, R. & PESOLE, G. 2013. ASPic-GeneID: A Lightweight Pipeline for Gene Prediction and Alternative Isoforms Detection. *BioMed research international,* 2013.

ALLEN, H. K., MOE, L. A., RODBUMRER, J., GAARDER, A. & HANDELSMAN, J. 2009. Functional metagenomics reveals diverse β-lactamases in a remote Alaskan soil. *The ISME journal,* 3**,** 243-251.

ALVAREZ, L. A., EXTON, D. A., TIMMIS, K. N., SUGGETT, D. J. & MCGENITY, T. J. 2009. Characterization of marine isoprene-degrading communities. *Environmental microbiology,* 11**,** 3280-3291.

ALVES L, MELOA M, MENDONC D, SIM˜OES F, MATOS J, TENREIRO R & F.M., G. 1. 2007. Sequencing, cloning and expression of the dsz genes required for dibenzothiophene sulfone desulfurization from Gordonia alkanivorans strain 1B. *Enzyme and microbial technology,* 40**,** 1598-1603.

AN, D., CAFFREY, S. M., SOH, J., AGRAWAL, A., BROWN, D., BUDWILL, K., DONG, X., DUNFIELD, P. F., FOGHT, J. & GIEG, L. M. 2013. Metagenomics of hydrocarbon resource environments indicates aerobic taxa and genes to be unexpectedly common. *Environmental science & technology,* 47**,** 10708-10717.

ANDERSON, D. C. & HAIRSTON, R. V. 1999. The Winogradsky Column & Biofilms: Models for Teaching Nutrient Cycling & Succession in an Ecosystem. *The American Biology Teacher***,** 453-459.

ANDERSON, K. W., BIERMANN, M., EIRICH, L. D., GATES, J. A., VICE, G. H. & WILSON, C. R. 2008. Biooxidation capabilities of Candida sp. Google Patents.

ANDOH, A., IMAEDA, H., AOMATSU, T., INATOMI, O., BAMBA, S., SASAKI, M., SAITO, Y., TSUJIKAWA, T. & FUJIYAMA, Y. 2011. Comparison of the fecal microbiota profiles between ulcerative colitis and Crohn's disease using terminal restriction fragment length polymorphism analysis. *Journal of gastroenterology,* 46**,** 479-486.

ANSORGE, W. J. 2009. Next-generation DNA sequencing techniques. *New biotechnology,* 25**,** 195-203.

ANTHONY, I. O. 2006. Biodegradation alternative in the cleanup of petroleum hydrocarbon pollutants. *Biotechnology and Molecular Biology Reviews,* 1**,** 38-50.

ARIAS, L. R. & YANG, L. 2009. Inactivation of bacterial pathogens by carbon nanotubes in suspensions. *Langmuir,* 25**,** 3003-3012.

ARIYARATNE, P. N. & SUNG, W.-K. 2011. PE-Assembler: de novo assembler using short paired-end reads. *Bioinformatics,* 27**,** 167-174.

ARORA, P. K., SRIVASTAVA, A. & SINGH, V. P. 2014. Bacterial degradation of nitrophenols and their derivatives. *Journal of hazardous materials,* 266**,** 42-59.

ASARI, M., OKA, K., OMURA, T., MASEDA, C., TASAKI, Y., SHIONO, H., MATSUBARA, K., MATSUDA, M. & SHIMIZU, K. 2013. Universal fluorescent labeling of amplification products using locked nucleic acids. *Electrophoresis,* 34**,** 448-455.

ASPERGER, O. & KLEHER, H. 1991. METABOLISM OF ALKANES BY AC IN ETOBACTER. *The Biology of Acinetobacter: Taxonomy, Clinical Importance, Molecular Biology, Physiology, Industrial Relevance,* 57**,** 323.

ATLAS, R. 1981. Microbial degradation of petroleum hydrocarbons: an environmental perspective. *Microbiology Review* 45**,** 180–209.

AUFFRET, M. D., YERGEAU, E., LABBÉ, D., FAYOLLE-GUICHARD, F. & GREER, C. W. 2014. Importance of Rhodococcus strains in a bacterial consortium degrading a mixture of hydrocarbons, gasoline, and diesel oil additives revealed by metatranscriptomic analysis. *Applied microbiology and biotechnology***,** 1-12.

B, V. D. 2005. The FadL family: unusual transporters for unusual substrates. *Current opinions in structural biology,* 15**,** 401-407.

BACCI, G., BAZZICALUPO, M., BENEDETTI, A. & MENGONI, A. 2013. StreamingTrim 1.0: a Java software for dynamic trimming of 16S rRNA sequence data from metagenetic studies. *Molecular ecology resources*.

BAHUGUNA, A., LILY, M. K., MUNJAL, A., SINGH, R. N. & DANGWAL, K. 2011. Desulfurization of dibenzothiophene (DBT) by a novel strain Lysinibacillus sphaericus DMT-7 isolated from diesel contaminated soil. *Journal of Environmental Sciences,* 23**,** 975-982.

BAKER, B. J., HUGENHOLTZ, P., DAWSON, S. C. & BANFIELD, J. F. 2003. Extremely acidophilic protists from acid mine drainage host Rickettsiales-lineage endosymbionts that have intervening sequences in their 16S rRNA genes. *Applied and environmental microbiology,* 69**,** 5512-5518.

BALBA, M., AL-AWADHI, N. & AL-DAHER, R. 1998a. Bioremediation of oil-contaminated soil: microbiological methods for feasibility assessment and field evaluation. *Journal of microbiological methods,* 32**,** 155-164.

BALBA, M., AL-DAHER, R., AL-AWADHI, N., CHINO, H. & TSUJI, H. 1998b. Bioremediation of oil-contaminated desert soil: the Kuwaiti experience. *Environment International,* 24**,** 163-173.

BARTLETT, D. Bacterial adaptation to extremes of low temperature and elevated pressure. 37th COSPAR Scientific Assembly, 2008. 198.

BASEN, M., KRÜGER, M., MILUCKA, J., KUEVER, J., KAHNT, J., GRUNDMANN, O., MEYERDIERKS, A., WIDDEL, F. & SHIMA, S. 2011. Bacterial enzymes for dissimilatory sulfate reduction in a marine microbial mat (Black Sea) mediating anaerobic oxidation of methane. *Environmental microbiology,* 13**,** 1370-1379.

BAZINET, A. L. & CUMMINGS, M. P. 2012. A comparative evaluation of sequence classification programs. *BMC bioinformatics,* 13**,** 92.

BEAL, E. J., HOUSE, C. H. & ORPHAN, V. J. 2009. Manganese- and iron-dependent marine methane oxidation. *Science,* 325**,** 184-187.

BEASLEY, K. K. & NANNY, M. A. 2012. Potential energy surface for anaerobic oxidation of methane via fumarate addition. *Environmental science & technology,* 46**,** 8244-8252.

BÉJA, O., ARAVIND, L., KOONIN, E. V., SUZUKI, M. T., HADD, A., NGUYEN, L. P., JOVANOVICH, S. B., GATES, C. M., FELDMAN, R. A. & SPUDICH, J. L. 2000. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science,* 289**,** 1902-1906.

BELL, C. E. Z. Geochemical aspects of the microbial modification of carbon compounds. Advances in organic geochemistry: proceedings of the international meeting in Milan, 1962, 2013. Elsevier, 339.

BELLER, H. R., GOH, E.-B. & KEASLING, J. D. 2010. Genes involved in long-chain alkene biosynthesis in Micrococcus luteus. *Applied and environmental microbiology,* 76**,** 1212-1223.

BEN AYED, H., JRIDI, M., MAALEJ, H., NASRI, M. & HMIDET, N. 2013. Characterization and stability of biosurfactant produced by Bacillus mojavensis A21 and its application in enhancing solubility of hydrocarbon. *Journal of Chemical Technology and Biotechnology*.

BERDUGO-CLAVIJO, C. & GIEG, L. M. 2014. Conversion of crude oil to methane by a microbial consortium enriched from oil reservoir production waters. *Frontiers in microbiology,* 5.

BEŠKOSKI, V. P., GOJGIĆ-CVIJOVIĆ, G., MILIĆ, J., ILIĆ, M., MILETIĆ, S., ŠOLEVIĆ, T. & VRVIĆ, M. M. 2011. bioremediation of a soil contaminated by mazut (heavy residual fuel oil)–A field experiment. *Chemosphere,* 83**,** 34-40.

BHATIA, S. & SHARMA, D. K. 2012. Thermophilic desulfurization of dibenzothiophene and different petroleum oils by Klebsiella sp. 13T. *Environmental Science and Pollution Research,* 19**,** 3491-3497.

BIŃ, A. K., MACHNIEWSKI, P., SAKOWICZ, R., OSTROWSKA, J. & ZIELIŃSKI, J. 2001. Degradation of Nitroaromatics (MNT, DNT and TNT) by AOPs. *OZONE SCIENCE AND ENGINEERING,* 23**,** 343-349.

BINAZADEH, M., KARIMI, I. A. & LI, Z. 2009. Fast biodegradation of long chain n-alkanes and crude oil at high concentrations with Rhodococcus sp. Moj-3449. *Enzyme and Microbial Technology,* 45**,** 195–202.

BIRNEY, E., CLAMP, M. & DURBIN, R. 2004. GeneWise and genomewise. *Genome research,* 14**,** 988-995.

BLANKENSHIP, R. E. 2013. *Molecular mechanisms of photosynthesis,* John Wiley & Sons.

BONHAM-CARTER, O., ALI, H. & BASTOLA, D. 2013. A base composition analysis of natural patterns for the preprocessing of metagenome sequences. *BMC Bioinformatics,* 14**,** S5.

BOONE, D. R., BRENNER, D. J., CASTENHOLZ, R. W., DE VOS, P., GARRITY, G. M., KRIEG, N. R. & STALEY, J. T. 2011. *Bergey's Manual of Systematic Bacteriology: The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes*, Springer.

BOOPATHY, R. 2000. Factors limiting bioremediation technologies. *Bioresource Technology,* 74**,** 63-67.

BOOPATHY, R., SHIELDS, S. & NUNNA, S. 2012. Biodegradation of crude oil from the BP oil spill in the marsh sediments of southeast Louisiana, USA. *Applied biochemistry and biotechnology,* 167**,** 1560-1568.

BORDOLOI, N. K., RAI, S. K., CHAUDHURI, M. K. & MUKHERJEE, A. K. 2014. Deep-desulfurization of dibenzothiophene and its derivatives present in diesel oil by a newly

isolated bacterium Achromobacter sp. to reduce the environmental pollution from fossil fuel combustion. *Fuel Processing Technology,* 119**,** 236-244.

BORONIN, A. & KOSHELEVA, I. 2010. Diversity of Naphthalene Biodegradation Systems in Soil Bacteria. *Handbook of Hydrocarbon and Lipid Microbiology.* Springer.

BORONIN, A. M. & KOSHELEVA, I. A. 2014. The Role of Catabolic Plasmids in Biodegradation of Petroleum Hydrocarbons. *Current Environmental Issues and Challenges.* Springer.

BOUWMAN, L., BEUSEN, A., GLIBERT, P. M., OVERBEEK, C., PAWLOWSKI, M., HERRERA, J., MULSOW, S., YU, R. & ZHOU, M. 2013. Mariculture: significant and expanding cause of coastal nutrient enrichment. *Environmental Research Letters,* 8**,** 044026.

BOYD, D. R., SHARMA, N. D., STEVENSON, P. J., BLAIN, M., MCROBERTS, C., HAMILTON, J. T., ARGUDO, J. M., MUNDI, H., KULAKOV, L. A. & ALLEN, C. C. 2011. Dioxygenase-catalysed cis-dihydroxylation of meta-substituted phenols to yield cyclohexenone cis-diol and derived enantiopure cis-triol metabolites. *Organic & biomolecular chemistry,* 9**,** 1479-1490.

BRADY, A. & SALZBERG, S. L. 2009. Phymm and PhymmBL: metagenomic phylogenetic classification with interpolated Markov models. *Nature methods,* 6**,** 673-676.

BRAKSTAD, O. G. & BONAUNET, K. 2006. Biodegradation of petroleum hydrocarbons in seawater at low temperatures (0–5 C) and bacterial communities associated with degradation. *Biodegradation,* 17**,** 71-82.

BROBERG, C. A. & CLARK, D. D. 2010. Shotgun proteomics of Xanthobacter autotrophicus Py2 reveals proteins specific to growth on propylene. *Archives of microbiology,* 192**,** 945-957.

BRODERSEN, J., BÄUMER, S., ABKEN, H. J., GOTTSCHALK, G. & DEPPENMEIER, U. 1999. Inhibition of membrane-bound electron transport of the methanogenic archaeon Methanosarcina mazei Gö1 by diphenyleneiodonium. *European journal of biochemistry / FEBS,* 259**,** 218.

BRODKORB, D., GOTTSCHALL, M., MARMULLA, R., LÜDDEKE, F. & HARDER, J. 2010a. Linalool dehydratase-isomerase, a bifunctional enzyme in the anaerobic degradation of monoterpenes. *The Journal of biological chemistry,* 285**,** 30436-30442.

BRODKORB, D., GOTTSCHALL, M., MARMULLA, R., LÜDDEKE, F. & HARDER, J. 2010b. Linalool dehydratase-isomerase@@@@, a bifunctional enzyme in the

anaerobic degradation of monoterpenes. *The Journal of biological chemistry,* 285**,** 30436-30442.

BRULC, J. M., ANTONOPOULOS, D. A., MILLER, M. E. B., WILSON, M. K., YANNARELL, A. C., DINSDALE, E. A., EDWARDS, R. E., FRANK, E. D., EMERSON, J. B. & WACKLIN, P. 2009. Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proceedings of the National Academy of Sciences,* 106**,** 1948-1953.

BUSCH, A., LACAL, J., SILVA-JÍMENEZ, H., KRELL, T. & RAMOS, J. L. 2010. Catabolite repression of the TodS/TodT two-component system and effector-dependent transphosphorylation of TodT as the basis for toluene dioxygenase catabolic pathway control. *Journal of bacteriology,* 192**,** 4246-4250.

BUSTAMANTE, M., DURÁN, N. & DIEZ, M. 2012. Biosurfactants are useful tools for the bioremediation of contaminated soil: a review. *Journal of soil science and plant nutrition,* 12**,** 667-687.

BUZANELLO, E. B., REZENDE, R. P., SOUSA, F. M., MARQUES, E. L. & LOGUERCIO, L. L. 2014. A novel Bacillus pumilus-related strain from tropical landfarm soil is capable of rapid dibenzothiophene degradation and biodesulfurization. *BMC microbiology,* 14**,** 257.

CAI, M., WANG, L., CAI, H., LI, Y., WANG, Y.-N., TANG, Y.-Q. & WU, X.-L. 2011. Salinarimonas ramus sp. nov. and Tessaracoccus oleiagri sp. nov., isolated from a crude oil-contaminated saline soil. *International journal of systematic and evolutionary microbiology,* 61**,** 1767-1775.

CALLAGHAN, A., MORRIS, B., PEREIRA, I., MCINERNEY, M., AUSTIN, R. N., GROVES, J. T., KUKOR, J., SUFLITA, J., YOUNG, L. & ZYLSTRA, G. 2012a. The genome sequence of Desulfatibacillum alkenivorans AK-01: a blueprint for anaerobic alkane oxidation. *Environmental microbiology,* 14**,** 101-113.

CALLAGHAN, A. V., MORRIS, B. E. L., PEREIRA, I. A. C., MCINERNEY, M. J., AUSTIN, R. N., GROVES, J. T., KUKOR, J. J., SUFLITA, J. M., YOUNG, L. Y., ZYLSTRA, G. J. & WAWRIK, B. 2012b. The genome sequence of Desulfatibacillum alkenivorans AK-01: a blueprint for anaerobic alkane oxidation. *Environmental Microbiology,* 14**,** 101-113.

CAPORASO, J. G., KUCZYNSKI, J., STOMBAUGH, J., BITTINGER, K., BUSHMAN, F. D., COSTELLO, E. K., FIERER, N., PENA, A. G., GOODRICH, J. K. & GORDON,

J. I. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature methods,* 7**,** 335-336.

CAPORASO, J. G., LAUBER, C. L., WALTERS, W. A., BERG-LYONS, D., LOZUPONE, C. A., TURNBAUGH, P. J., FIERER, N. & KNIGHT, R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences,* 108**,** 4516-4522.

CAPPELLETTI, M., FEDI, S., FRASCARI, D., OHTAKE, H., TURNER, R. & ZANNONI, D. 2011. Analyses of both the alkB gene transcriptional start site and alkB promoter-inducing properties of Rhodococcus sp. strain BCP1 grown on n-alkanes. *Applied and environmental microbiology,* 77**,** 1619-1627.

CARDOSO, A. M. & COUTINHO, F. H. 2012a. Metagenomics in polluted @@@aquatic environments.

CARDOSO, A. M. & COUTINHO, F. H. 2012b. Metagenomics in polluted aquatic environments.

CARTEGNI, L., WANG, J., ZHU, Z., ZHANG, M. Q. & KRAINER, A. R. 2003. ESEfinder: a web resource to identify exonic splicing enhancers. *Nucleic acids research,* 31**,** 3568-3571.

CAWLEY, S. 2000. The Genscan HMM.

CECCHINI, D. A., LAVILLE, E., LAGUERRE, S., ROBE, P., LECLERC, M., DORÉ, J., HENRISSAT, B., REMAUD-SIMÉON, M., MONSAN, P. & POTOCKI-VÉRONÈSE, G. 2013. Functional Metagenomics Reveals Novel Pathways of Prebiotic Breakdown by Human Gut Bacteria. *PloS one,* 8**,** e72766.

CHANG, J. I. & LIN, C.-C. 2006. A study of storage tank accidents. *Journal of loss prevention in the process industries,* 19**,** 51-59.

CHANG, J. Y., ANTONOPOULOS, D. A., KALRA, A., TONELLI, A., KHALIFE, W. T., SCHMIDT, T. M. & YOUNG, V. B. 2008. Decreased Diversity of the Fecal Microbiome in Recurrent Clostridium difficile—Associated Diarrhea. *Journal of Infectious Diseases,* 197**,** 435-438.

CHATTOPADHYAY, M. 2006. Mechanism of bacterial adaptation to low temperature. *Journal of biosciences,* 31**,** 157-165.

CHAUHAN, A. K., SINGH, S. P. & KUMAR, A. 2014. Cloning, Sequencing and Characterization of the 'Desulfurization Gene Cluster'From an Isolated Strain Gordonia sp. IITR100. *Journal of Applied Biotechnology,* 2**,** p65.

CHEN, C., KHALEEL, S. S., HUANG, H. & WU, C. H. ngsShoRT: A Software for Pre-processing Illumina Short Read Sequences for De Novo Genome Assembly. Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics, 2013. ACM, 706.

CHEN, M. W., CHENG, Q., GIBSON, K. J., KOSTICHKA, K. N.-S., THOMAS, S. M. & NAGARAJAN, V. 2004. Genes involved in cyclododecanone degradation pathway. Google Patents.

CHEN, Y., LI, C., ZHOU, Z., WEN, J., YOU, X., MAO, Y., LU, C., HUO, G. & JIA, X. 2014. Enhanced Biodegradation of Alkane Hydrocarbons and Crude Oil by Mixed Strains and Bacterial Community Analysis. *Applied biochemistry and biotechnology,* 172**,** 3433-3447.

CHENG, Q., THOMAS, S. & ROUVIERE, P. 2002. Biological conversion of cyclic alkanes and cyclic alcohols into dicarboxylic acids: biochemical and molecular basis. *Applied microbiology and biotechnology,* 58**,** 704-711.

CHOI, E. N., CHO, M. C., KIM, Y., KIM, C.-K. & LEE, K. 2003. Expansion of growth substrate range in Pseudomonas putida F1 by mutations in both cymR and todS, which recruit a ring-fission hydrolase CmtE and induce the tod catabolic operon, respectively. *Microbiology,* 149**,** 795-805.

CHOU, C.-C., RIVIERE, J. E. & MONTEIRO-RIVIERE, N. A. 2002. Differential relationship between the carbon chain length of jet fuel aliphatic hydrocarbons and their ability to induce cytotoxicity vs. interleukin-8 release in human epidermal keratinocytes. *Toxicological Sciences,* 69**,** 226-233.

CHUNSHUANG, L., DONGFENG, Z. & YUNBO, Z. 2014. Effects of temperature, acetate and nitrate on methane generation from petroleum hydrocarbons [J]. *China Petroleum Processing and Petrochemical Technology,* 16**,** 24-31.

CHURCH, G. M., VIGNEAULT, F. & SISMOUR, A. M. 2011. ENZYMATIC OLIGONUCLEOTIDE PRE-ADENYLATION. Google Patents.

CLARK, D. D. & ENSIGN, S. A. 1999. Evidence for an Inducible Nucleotide-Dependent Acetone Carboxylase in Rhodococcus rhodochrousB276. *Journal of bacteriology,* 181**,** 2752-2758.

COATES, J. D., CHAKRABORTY, R., LACK, J. G., O'CONNOR, S. M., COLE, K. A., BENDER, K. S. & ACHENBACH, L. A. 2001. Anaerobic benzene oxidation coupled to nitrate reduction in pure culture by two strains of Dechloromonas. *Nature,* 411**,** 1039-1043.

COLEMAN, N. V., LE, N. B., LY, M. A., OGAWA, H. E., MCCARL, V., WILSON, N. L. & HOLMES, A. J. 2011a. Hydrocarbon monooxygenase in Mycobacterium: recombinant expression of a member of the ammonia monooxygenase superfamily. *The ISME journal,* 6**,** 171-182.

COLEMAN, N. V., YAU, S., WILSON, N. L., NOLAN, L. M., MIGOCKI, M. D., LY, M. A., CROSSETT, B. & HOLMES, A. J. 2011b. Untangling the multiple monooxygenases of Mycobacterium chubuense strain NBB4, a versatile hydrocarbon degrader. *Environmental microbiology reports,* 3**,** 297-307.

COLLIER, T. K., CHIANG, M., AU, D. & RAINBOW, P. S. 2012. Biomarkers Currently Used in Environmental Monitoring. *Ecological Biomarkers: Indicators of Ecotoxicological Effects, CRC Press, Boca Raton***,** 385-410.

COLOMBO, M., CAVALCA, L., BERNASCONI, S. & ANDREONI, V. 2011. Bioremediation of polyaromatic hydrocarbon contaminated soils by native microflora and bioaugmentation with Sphingobium chlorophenolicum strain C3R: A feasibility study in solid-and slurry-phase microcosms. *International Biodeterioration & Biodegradation,* 65**,** 191-197.

COMANDINI, A., DUBOIS, T., ABID, S. & CHAUMEIX, N. 2013. Comparative study on cyclohexane and decalin oxidation. *Energy & Fuels,* 28**,** 714-724.

CONANT, R. T., RYAN, M. G., ÅGREN, G. I., BIRGE, H. E., DAVIDSON, E. A., ELIASSON, P. E., EVANS, S. E., FREY, S. D., GIARDINA, C. P. & HOPKINS, F. M. 2011. Temperature and soil organic matter decomposition rates–synthesis of current knowledge and a way forward. *Global Change Biology,* 17**,** 3392-3404.

CONG, Q. & GRISHIN, N. V. 2012. MESSA: MEta-server for protein sequence analysis. *BMC biology,* 10**,** 82.

CONSORTIUM, H. M. P. 2012. Structure, function and diversity of the healthy human microbiome. *Nature,* 486**,** 207-214.

CORTES, P., DENG, S. & SMITH, G. B. 2014. The Toxic Effects of Single Wall Carbon Nanotubes on E. coli and a Spore-Forming Bacillus Species. *Nanoscience and Nanotechnology Letters,* 6**,** 26-30.

COX, M. P., PETERSON, D. A. & BIGGS, P. J. 2010. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC bioinformatics,* 11**,** 485.

CRAVO-LAUREAU, C., GROSSI, V., RAPHEL, D., MATHERON, R. & HIRSCHLER-RÉA, A. 2005. Anaerobic n-alkane metabolism by a sulfate-reducing bacterium,

Desulfatibacillum aliphaticivorans strain CV2803T. *Applied and environmental microbiology,* 71**,** 3458-3467.

ČREŠNAR, B. & PETRIČ, Š. 2011. Cytochrome P450 enzymes in the fungal kingdom. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics,* 1814**,** 29-35.

CRISCUOLO, A. & BRISSE, S. 2013. AlienTrimmer: A tool to quickly and accurately trim off multiple short contaminant sequences from high-throughput sequencing reads. *Genomics,* 102**,** 500-506.

CRYAN, J. F. & DINAN, T. G. 2012. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nature reviews neuroscience,* 13**,** 701-712.

CUBITTO, M. A., MORAN, A. C., COMMENDATORE, M., CHIARELLO, M. N., BALDINI, M. D. & SINERIZ, F. 2004. Effects of Bacillus subtilis O9 biosurfactant on the bioremediation of crude oil-polluted soils. *Biodegradation,* 15**,** 281-287.

DAI, M. & COPLEY, S. D. 2004. Genome shuffling improves degradation of the anthropogenic pesticide pentachlorophenol by Sphingobium chlorophenolicum ATCC 39723. *Applied and environmental microbiology,* 70**,** 2391-2397.

DAI, Y., SHAO, R., QI, G. & DING, B.-B. 2014. Enhanced dibenzothiophene biodesulfurization by immobilized cells of Brevibacterium lutescens in n-octane–water biphasic system. *Applied biochemistry and biotechnology,* 174**,** 2236-2244.

DALHUS, B., SAARINEN, M., SAUER, U. H., EKLUND, P., JOHANSSON, K., KARLSSON, A., RAMASWAMY, S., BJORK, A., SYNSTAD, B. & NATERSTAD, K. 2002. Structural basis for thermophilic protein stability: structures of thermophilic and mesophilic malate dehydrogenases. *Journal of molecular biology,* 318**,** 707-721.

DANKO, A. S. & FREEDMAN, D. L. 2008. Involvement of carbon dioxide in the aerobic biodegradation of ethylene oxide, ethene, and vinyl chloride. *Process Biochemistry,* 43**,** 517-521.

DAVENPORT, C. F. & TÜMMLER, B. 2013. Advances in computational analysis of metagenome sequences. *Environmental microbiology,* 15**,** 1-5.

DAYAL, P. V., SINGH, H., BUSENLEHNER, L. & ELLIS, H. 2013. Identifying the Mechanism of Reduced Flavin Transfer in Alkanesulfonate Monooxygenase System. *The FASEB Journal,* 27**,** 789.21.

DE LA TORRE, J. R., WALKER, C. B., INGALLS, A. E., KÖNNEKE, M. & STAHL, D. A. 2008. Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. *Environmental Microbiology,* 10**,** 810-818.

DE MARVAO, A., DAWES, T., KEENAN, N. G., BAI, W., SHI, W., DURIGHEL, G., DIAMOND, T., MONJE-GARCIA, L., COOK, S. A. & O'REGAN, D. P. 2013. The UK GenScan study-population-based imaging genetics research using 3D Cardiac Magnetic Resonance. *Journal of Cardiovascular Magnetic Resonance,* 15**,** E2.

DEL FABBRO, C., SCALABRIN, S., MORGANTE, M. & GIORGI, F. M. 2013. An Extensive Evaluation of Read Trimming Effects on Illumina NGS Data Analysis. *PloS one,* 8**,** e85024.

DELA TORRE, R., LARKIN, J., SINGER, A. & MELLER, A. 2012. Fabrication and characterization of solid-state nanopore arrays for high-throughput DNA sequencing. *Nanotechnology,* 23**,** 385308.

DELCHER, A. L., HARMON, D., KASIF, S., WHITE, O. & SALZBERG, S. L. 1999. Improved microbial gene identification with GLIMMER. *Nucleic acids research,* 27**,** 4636-4641.

DELLA SALA, G., HOCHMUTH, T., COSTANTINO, V., TETA, R., GERWICK, W., GERWICK, L., PIEL, J. & MANGONI, A. 2013. Polyketide genes in the marine sponge Plakortis simplex: a new group of mono-modular type I polyketide synthases from sponge symbionts. *Environmental microbiology reports,* 5**,** 809-818.

DELSENY, M., HAN, B. & HSING, Y. I. 2010. High throughput DNA sequencing: the new sequencing revolution. *Plant Science,* 179**,** 407-422.

DENEF, V. J., KALNEJAIS, L. H., MUELLER, R. S., WILMES, P., BAKER, B. J., THOMAS, B. C., VERBERKMOES, N. C., HETTICH, R. L. & BANFIELD, J. F. 2010. Proteogenomic basis for ecological divergence of closely related bacteria in natural acidophilic microbial communities. *Proceedings of the National Academy of Sciences,* 107**,** 2383-2390.

DESAI, N., ANTONOPOULOS, D., GILBERT, J. A., GLASS, E. M. & MEYER, F. 2012. From genomics to metagenomics. *Current Opinion in Biotechnology,* 23**,** 72-76.

DESMET, F.-O., HAMROUN, D., LALANDE, M., COLLOD-BÉROUD, G., CLAUSTRES, M. & BÉROUD, C. 2009. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic acids research,* 37**,** e67-e67.

DIAZ, N. N., KRAUSE, L., GOESMANN, A., NIEHAUS, K. & NATTKEMPER, T. W. 2009. TACOA–Taxonomic classification of environmental genomic fragments using a kernelized nearest neighbor approach. *BMC bioinformatics,* 10**,** 56.

DINAMARCA, M. A., ARANDA-OLMEDO, I., PUYET, A. & ROJO, F. 2003. Expression of the Pseudomonas putida OCT plasmid alkane degradation pathway is modulated by

two different global control signals: evidence from continuous cultures. *Journal of bacteriology,* 185**,** 4772-4778.

DIONISI, H. M., LOZADA, M. & OLIVERA, N. L. 2012. Bioprospection of marine microorganisms: biotechnological applications and methods. *Revista argentina de microbiología,* 44**,** 49-60.

DOBSLAW, D. & ENGESSER, K. H. 2014. Degradation of toluene by ortho cleavage enzymes in Burkholderia fungorum FLU100. *Microbial biotechnology*.

DOHM, J. C., LOTTAZ, C., BORODINA, T. & HIMMELBAUER, H. 2007. SHARCGS, a fast and highly accurate short-read assembly algorithm for de novo genomic sequencing. *Genome research,* 17**,** 1697-1706.

DOJKA, M. A., HUGENHOLTZ, P., HAACK, S. K. & PACE, N. R. 1998. Microbial Diversity in a Hydrocarbon- and Chlorinated-Solvent-Contaminated Aquifer Undergoing Intrinsic Bioremediation. *Applied and environmental Microbiology,* 64**,** 3869–3877.

DONG, X., TANG, Y., WU, M., VLAHOVIC, B. & YANG, L. 2013. Dual effects of single-walled carbon nanotubes coupled with near-infrared radiation on Bacillus anthracis spores: inactivates spores and stimulates the germination of surviving spores. *Journal of biological engineering,* 7**,** 19.

DORN, P. B., VIPOND, T. E., SALANITRO, J. P. & WISNIEWSKI, H. L. 1998. Assessment of the acute toxicity of crude oils in soils using earthworms, microtox and plants. *Chemosphere,* 37**,** 845-860.

DOS SANTOS, V. M., YAKIMOV, M. M., TIMMIS, K. N. & GOLYSHIN, P. N. 2008. Genomic insights into oil biodegradation in marine systems. *Microbial biodegradation: genomic and molecular biology*.

DUTTA, S., THATOI, H., BEHERA, B. & MISHRA, R. 2014. Sulphur oxidising bacteria in mangrove ecosystem: A review. *African Journal of Biotechnology,* 13**,** 2897-2907.

DWORKIN, M. & GUTNICK, D. 2012. Sergei Winogradsky: a founder of modern microbiology and the first microbial ecologist. *FEMS microbiology reviews,* 36**,** 364-379.

EDWARDS, K. J., BOND, P. L., DRUSCHEL, G. K., MCGUIRE, M. M., HAMERS, R. J. & BANFIELD, J. F. 2000. Geochemical and biological aspects of sulfide mineral dissolution: lessons from Iron Mountain, California. *Chemical Geology,* 169**,** 383-397.

EFROYMSON, R. A. & ALEXANDER, M. 1991. Biodegradation by an Arthrobacter species of hydrocarbons partitioned into an organic solvent. *Applied and environmental microbiology,* 57**,** 1441-1447.

EHRENREICH, P., BEHRENDS, A., HARDER, J. & WIDDEL, F. 2000. Anaerobic oxidation of alkanes by newly isolated denitrifying bacteria. *Archives of microbiology,* 173**,** 58-64.

EINSLE, O., NIESSEN, H., ABT, D. J., SEIFFERT, G., SCHINK, B., HUBER, R., MESSERSCHMIDT, A. & KRONECK, P. M. 2005. Crystallization and preliminary X-ray analysis of the tungsten-dependent acetylene hydratase from Pelobacter acetylenicus. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications,* 61**,** 299-301.

ENGEL, A. S. & GUPTA, A. A. 2014. Regime Shift in Sandy Beach Microbial Communities following Deepwater Horizon Oil Spill Remediation Efforts. *PloS one,* 9**,** e102934.

ENNING, D., VENZLAFF, H., GARRELFS, J., DINH, H. T., MEYER, V., MAYRHOFER, K., HASSEL, A. W., STRATMANN, M. & WIDDEL, F. 2012. Marine sulfate-reducing bacteria cause serious corrosion of iron under electroconductive biogenic mineral crust. *Environmental microbiology,* 14**,** 1772-1787.

ERB, T. J. 2011. Carboxylases in natural and synthetic microbial pathways. *Applied and environmental microbiology,* 77**,** 8466-8477.

ERICKSON, R., MYERS, A. T. & HORR, C. A. 1954. Association of uranium and other metals with crude oil, asphalt, and petroliferous rock. *AAPG Bulletin,* 38**,** 2200-2218.

ESCHENFELDT, W. H., STOLS, L., ROSENBAUM, H., KHAMBATTA, Z. S., QUAITE-RANDALL, E., WU, S., KILGORE, D. C., TRENT, J. D. & DONNELLY, M. I. 2001. DNA from uncultured organisms as a source of 2, 5-diketo-D-gluconic acid reductases. *Applied and environmental microbiology,* 67**,** 4206-4214.

ETTWIG, K. F., SPETH, D. R., REIMANN, J., WU, M. L., JETTEN, M. S. M. & KELTJENS, J. T. 2012. Bacterial oxygen production in the dark. *Frontiers in microbiology,* 3**,** 273.

EVANS, F. F., ROSADO, A. S., SEBASTIÁN, G. V., CASELLA, R., MACHADO, P. L., HOLMSTRÖM, C., KJELLEBERG, S., VAN ELSAS, J. D. & SELDIN, L. 2004. Impact of oil contamination and biostimulation on the diversity of indigenous bacterial communities in soil microcosms. *FEMS Microbiology Ecology,* 49**,** 295-305.

FAHEEM, M., VIDAL, J., FERNANDES, J., OLIVEIRA, G., CAROLINE, C., SOUTO, B., PARACHIN, N., QUIRINO, B. & BARBOSA, J. Cloning and Initial Characterization of Cellulolytic Enzymes Selected in a Metagenomic Approach. Embrapa Agroenergia-Resumo em anais de congresso (ALICE), 2013. In: REUNIÃO ANUAL DA SBBQ, 42., 2013, Foz do Iguaçu, PR.[Anais...] São Paulo: SBBq, 2013.

FAKRUDDIN, R. M. M., CHOWDHURY, A., HOSSAIN, N., MAHAJAN, S. & ISLAM, S. 2013. Pyrosequencing-A Next Generation Sequencing Technology. *World Applied Sciences Journal,* 24**,** 1558-1571.

FAVIER, C. F., DE VOS, W. M. & AKKERMANS, A. D. 2003. Development of bacterial and bifidobacterial communities in feces of newborn babies. *Anaerobe,* 9**,** 219-229.

FELSENSTEIN, J. 2006. PHYLIP (Phylogeny Inference Package) documentation files, version 3.66. *Seattle: Department of Genome Sciences, University of Washington*.

FENG, Y., KHOO, H. E. & POH, C. L. 1999. Purification and characterization of gentisate 1, 2-dioxygenases from Pseudomonas alcaligenes NCIB 9867 and Pseudomonas putida NCIB 9869. *Applied and environmental microbiology,* 65**,** 946-950.

FERRARO, D. J., GAKHAR, L. & RAMASWAMY, S. 2005. Rieske business: structure–function of Rieske non-heme oxygenases. *Biochemical and biophysical research communications,* 338**,** 175-190.

FINEGOLD, S. M. 2008. Therapy and epidemiology of autism–clostridial spores as key elements. *Medical hypotheses,* 70**,** 508-511.

FINGAS, M. 2010. *Oil spill science and technology,* Gulf professional publishing.

FINGAS, M. 2012. *The basics of oil spill cleanup*, CRC Press.

FIROUZ ABBASIAN, ROBIN LOCKINGTON, MALLAVARAPU MEGHARAJ & NAIDU, R. 2015. A Review on the Genetics of Aliphatic and Aromatic Hydrocarbon Degradation. *Applied biochemistry and biotechnology,* In press**,** 1-27.

FIROUZ ABBASIAN, ROBIN LOCKINGTON, THAVAMANI PALANISAMI, KAVITHA RAMADASS, MALLAVARAPU MEGHARAJ & NAIDU, R. 2016. Microbial diversity and hydrocarbon degrading gene capacity of a crude oil field soil as determined by metagenomic analysis. *Biotechnology Progress,* In Press.

FONDI, M., RIZZI, E., EMILIANI, G., ORLANDINI, V., BERNA, L., PAPALEO, M. C., PERRIN, E., MAIDA, I., CORTI, G. & DE BELLIS, G. 2013. The genome sequence of the hydrocarbon-degrading Acinetobacter venetianus VE-C3. *Research in microbiology,* 164**,** 439-449.

FUHRMAN, J. A., CRAM, J. A. & NEEDHAM, D. M. 2015. Marine microbial community dynamics and their ecological interpretation. *Nature Reviews Microbiology*.

FUNK, S., CRAWFORD, D., CRAWFORD, R., MEAD, G. & DAVIS-HOOVER, W. 1995. Full-scale anaerobic bioremediation of trinitrotoluene (TNT) contaminated soil. *Applied Biochemistry and Biotechnology,* 51**,** 625-633.

FURUKAWA, K., SUENAGA, H. & GOTO, M. 2004. Biphenyl dioxygenases: functional versatilities and directed evolution. *Journal of bacteriology*, 186**,** 5189-5196.

GABRIEL CORREA-REYES, J., PILAR SANCHEZ-SAAVEDRA, M. D., ALFARO SIQUEIROS-BELTRONES, D. & FLORES-ACEVEDO, N. 2001. Isolation and growth of eight strains of benthic diatoms, cultured under two light conditions. *Journal of Shellfish Research*, 20**,** 603-610.

GALLEGO, J. L., LOREDO, J., LLAMAS, J. F., VÁZQUEZ, F. & SÁNCHEZ, J. 2001. Bioremediation of diesel-contaminated soils: Evaluation of potential in situ techniques by study of bacterial degradation. *Biodegradation,* 12**,** 325-335.

GALLEGOS, M.-T., MOLINA-HENARES, A. J., ZHANG, X., TERÁN, W., BERNAL, P., ALGUEL, Y., GUAZZARONI, M.-E., KRELL, T., SEGURA, A. & RAMOS, J.-L. 2008. Genomic insights into solvent tolerance and pumps that extrude toxic chemicals. *Microbial Biodegradation: Genomics and Molecular Biology*.

GANZERT, L., LIPSKI, A., HUBBERTEN, H.-W. & WAGNER, D. 2011. The impact of different soil parameters on the community structure of dominant bacteria from nine different soils located on Livingston Island, South Shetland Archipelago, Antarctica. *FEMS microbiology ecology*, 76**,** 476-491.

GARCIA-ARMISEN, T., VERCAMMEN, K., PASSERAT, J., TRIEST, D., SERVAIS, P. & CORNELIS, P. 2011. Antimicrobial resistance of heterotrophic bacteria in sewage-contaminated rivers. *Water research*, 45**,** 788-796.

GARCIA-MENDOZA, E. & OCAMPO-ALVAREZ, H. 2011. Photoprotection in the brown alga Macrocystis pyrifera: evolutionary implications. *Journal of Photochemistry and Photobiology B: Biology*, 104**,** 377-385.

GARCIA-PICHEL, F. & CASTENHOLZ, R. W. 1991. Characterization and biological implications if Scytonemin, a cyanobacterial shealth pigment. *Journal of Phycology*, 27**,** 395-409.

GARGOURI, B., KARRAY, F., MHIRI, N., ALOUI, F. & SAYADI, S. 2013. Bioremediation of petroleum hydrocarbons-contaminated soil by bacterial consortium isolated from an industrial wastewater treatment plant. *Journal of Chemical Technology and Biotechnology*.

GARMENDIA, L., HERNANDEZ, A., SANCHEZ, M. & MARTINEZ, J. 2012. Metagenomics and antibiotics. *Clinical Microbiology and Infection,* 18**,** 27-31.

GARRITY, G., BELL, J. & LILBURN, T. 2005. Order I. Rhodospirillales Pfennig and Trüper 1971, 17 AL. *Bergey's Manual of Systematic Bacteriology***,** 1-95.

GARRITY, G. M. & HOLT, J. G. 2001. Phylum BVIII. Nitrospirae phy. nov. *Bergey's Manual® of Systematic Bacteriology*. Springer.

GENTILI, A. R., CUBITTO, M. A., FERRERO, M. & RODRIGUÉZ, M. S. 2006. Bioremediation of crude oil polluted seawater by a hydrocarbon-degrading bacterial strain immobilized on chitin and chitosan flakes. *International Biodeterioration & Biodegradation,* 57**,** 222-228.

GERDES, B., BRINKMEYER, R., DIECKMANN, G. & HELMKE, E. 2005. Influence of crude oil on changes of bacterial communities in Arctic sea-ice. *FEMS Microbiology Ecology,* 53**,** 129-139.

GHANEM, K. M., EL-AASSAR, S. A. & YUSEF, H. H. 1992. Transformation of Reichstein's compound S into prednisolone by immobilized mixed cultures. *Journal of Chemical Technology and Biotechnology,* 54**,** 115-121.

GHAZALI, F. M., RAHMAN, R. N. Z. A., SALLEH, A. B. & BASRI, M. 2004. Biodegradation of hydrocarbons in soil by microbial consortium. *International Biodeterioration & Biodegradation,* 54**,** 61-67.

GIBSON, D. T. & PARALES, R. E. 2000. Aromatic hydrocarbon dioxygenases in environmental biotechnology. *Current opinion in biotechnology,* 11**,** 236-243.

GICH, F. & OVERMANN, J. 2006. Sandarakinorhabdus limnophila gen. nov., sp. nov., a novel bacteriochlorophyll a-containing, obligately aerobic bacterium isolated from freshwater lakes. *International journal of systematic and evolutionary microbiology,* 56**,** 847-854.

GIEG, L. 2010. How Specific Microbial Communities Benefit the Oil Industry: Case Study–Proof of Concept that Oil Entrained in Marginal Reservoirs Can Be Bioconverted to Methane Gas as a Green Energy Recovery Strategy. *Applied Microbiology and Molecular Biology in Oilfield Systems.* Springer.

GIEG, L. M., DUNCAN, K. E. & SUFLITA, J. M. 2008. Bioenergy production via microbial conversion of residual oil to natural gas. *Applied and environmental microbiology,* 74**,** 3022-3029.

GIRHARD, M., KLAUS, T., KHATRI, Y., BERNHARDT, R. & URLACHER, V. B. 2010. Characterization of the versatile monooxygenase CYP109B1 from Bacillus subtilis. *Applied microbiology and biotechnology,* 87**,** 595-607.

GLASS, E. M. & MEYER, F. 2011. The Metagenomics RAST Server: A Public Resource for the Automatic Phylogenetic and Functional Analysis of Metagenomes. *Handbook of*

*Molecular Microbial Ecology I: Metagenomics and Complementary Approaches***,** 325-331.

GLASS, E. M., WILKENING, J., WILKE, A., ANTONOPOULOS, D. & MEYER, F. 2010. Using the metagenomics RAST server (MG-RAST) for analyzing shotgun metagenomes. *Cold Spring Harbor Protocols,* 2010**,** pdb. prot5368.

GLUECK, S. M., GÜMÜS, S., FABIAN, W. M. & FABER, K. 2010. Biocatalytic carboxylation. *Chemical Society Reviews,* 39**,** 313-328.

GOETZ, F. E. & HARMUTH, L. J. 1992. Gentisate pathway in Salmonella typhimurium: metabolism of m-hydroxybenzoate and gentisate. *FEMS microbiology letters,* 97**,** 45-49.

GOGOI, B., DUTTA, N., GOSWAMI, P. & KRISHNA MOHAN, T. 2003. A case study of bioremediation of petroleum-hydrocarbon contaminated soil at a crude oil spill site. *Advances in Environmental Research,* 7**,** 767-782.

GOŁĘBIEWSKI, M., DEJA-SIKORA, E., CICHOSZ, M., TRETYN, A. & WRÓBEL, B. 2014. 16S rDNA pyrosequencing analysis of bacterial community in heavy metals polluted soils. *Microbial ecology,* 67**,** 635-647.

GONDAL, M., HUSSAIN, T., YAMANI, Z. & BAIG, M. 2006. Detection of heavy metals in Arabian crude oil residue using laser induced breakdown spectroscopy. *Talanta,* 69**,** 1072-1078.

GONG, J.-S., LU, Z.-M., LI, H., ZHOU, Z.-M., SHI, J.-S. & XU, Z.-H. 2013. Metagenomic technology and genome mining: emerging areas for exploring novel nitrilases. *Applied microbiology and biotechnology,* 97**,** 6603-6611.

GOODFELLOW, M., PETER, K., BUSSE, H.-J. R., TRUJILLO, M. E., LUDWIG, W., SUZUKI, K.-I. & PARTE, A. 2012. *Bergey's Manual of Systematic Bacteriology: Volume 5: The Actinobacteria,* Springer New York.

GOSALBES, M. J., DURBÁN, A., PIGNATELLI, M., ABELLAN, J. J., JIMÉNEZ-HERNÁNDEZ, N., PÉREZ-COBAS, A. E., LATORRE, A. & MOYA, A. 2011. Metatranscriptomic approach to analyze the functional human gut microbiota. *PloS one,* 6**,** e17447.

GOYAL, A. & ZYLSTRA, G. 1997. Genetics of naphthalene and phenanthrene degradation by Comamonas testosteroni. *Journal of Industrial Microbiology and Biotechnology,* 19**,** 401-407.

GRABOWSKI, A., NERCESSIAN, O., FAYOLLE, F., BLANCHET, D. & JEANTHON, C. 2005. Microbial diversity in production waters of a low-temperature biodegraded oil reservoir. *FEMS Microbiology Ecology,* 54**,** 427-443.

GREEN, B. & PARSON, W. W. 2003. *Light-harvesting antennas in photosynthesis*, Springer Science & Business Media.

GREEN, S., LEIGH, M. & NEUFELD, J. 2010. Denaturing gradient gel electrophoresis (DGGE) for microbial community analysis. *Handbook of Hydrocarbon and Lipid Microbiology.* Springer.

GRICMAN, Ł., VOGEL, C. & PLEISS, J. 2013. Conservation analysis of class-specific positions in cytochrome P450 monooxygenases: Functional and structural relevance. *Proteins: Structure, Function, and Bioinformatics*.

GROSSI, V., CRAVO-LAUREAU, C., GUYONEAUD, R., RANCHOU-PEYRUSE, A. & HIRSCHLER-RÉA, A. 2008. Metabolism of n-alkanes and n-alkenes by anaerobic bacteria: A summary. *Organic Geochemistry,* 39**,** 1197-1203.

GROSSI, V., CRAVO-LAUREAU, C., MÉOU, A., RAPHEL, D., GARZINO, F. & HIRSCHLER-RÉA, A. 2007. Anaerobic 1-alkene metabolism by the alkane- and alkene-degrading sulfate reducer Desulfatibacillum aliphaticivorans strain CV2803T. *Applied and environmental microbiology,* 73**,** 7882-7890.

GROSSI, V., CRAVO-LAUREAU, C., RONTANI, J.-F., CROS, M. & HIRSCHLER-RÉA, A. 2011. Anaerobic oxidation of n-alkenes by sulphate-reducing bacteria from the genus Desulfatiferula: n-ketones as potential metabolites. *Research in microbiology,* 162**,** 915.

GUENGERICH, F. P. & MUNRO, A. W. 2013. Unusual cytochrome P450 enzymes and reactions. *Journal of Biological Chemistry,* 288**,** 17065-17073.

GUERMOUCHE, M., BENSALAH, F. & GRAY, N. 2013. Application of Molecular Methods as a Biomarker in Bioremediation Studies. *Inter-national Journal of Biotechnology Applications, ISSN***,** 0975-2943.

GUHA, P. & GUPTA, A. 2013. RECOVERY OF GLYCERINE FROM SOAP-LYE. *Journal of the Indian Institute of Science,* 23**,** 317.

GUO, X., DING, X., MENG, Y. & PAN, Y. 2013. Cloud Computing for De Novo Metagenomic Sequence Assembly. *Bioinformatics Research and Applications.* Springer.

GUPTA, S., MARKO, M. G., MILLER, V. A., SCHAEFER, F. T., ANTHONY, J. R. & PORTER, J. R. 2014. Novel Production of Terpenoids in Escherichia coli and Activities Against Breast Cancer Cell Lines. *Applied biochemistry and biotechnology***,** 1-13.

GUTIERREZ, T. 2011. Identifying polycyclic aromatic hydrocarbon-degrading bacteria in oil-contaminated surface waters at Deepwater Horizon by cultivation, stable isotope probing and pyrosequencing. *Reviews in Environmental Science and Bio/Technology,* 10**,** 301-305.

GW, T., J, C., P, H., EE, A., RJ, R., PM, R., VV, S., EM, R., DS, R. & JF., B. 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature,* 428**,** 37-43.

HABE, H. & OMORI, T. 2003. Genetics of polycyclic aromatic hydrocarbon metabolism in diverse aerobic bacteria. *Bioscience, biotechnology, and biochemistry,* 67**,** 225-243.

HAIGLER, B. E., NISHINO, S. F. & SPAIN, J. C. 1994. Biodegradation of 4-methyl-5-nitrocatechol by Pseudomonas sp. strain DNT. *Journal of bacteriology,* 176**,** 3433-3437.

HAISER, H. J. & TURNBAUGH, P. J. 2013. Developing a metagenomic view of xenobiotic metabolism. *Pharmacological Research,* 69**,** 21-31.

HALECKY, M., KARLOVA, P., PACA, J., STIBOROVA, M., KOZLIAK, E. I., BAJPAI, R. & SEDLACEK, I. 2013. Biodegradation of a mixture of mononitrophenols in a packed-bed aerobic reactor. *Journal of Environmental Science and Health, Part A,* 48**,** 989-999.

HAMADY, M. & KNIGHT, R. 2009. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome research,* 19**,** 1141-1152.

HAMELIN, R. C., BÉRUBÉ, P., GIGNAC, M. & BOURASSA, M. 1996. Identification of root rot fungi in nursery seedlings by nested multiplex PCR. *Applied and Environmental Microbiology,* 62**,** 4026-4031.

HAN, D., PAL, S., YANG, Y., JIANG, S., NANGREAVE, J., LIU, Y. & YAN, H. 2013. DNA gridiron nanostructures based on four-arm junctions. *Science,* 339**,** 1412-1415.

HAO, R., LU, A. & WANG, G. 2004. Crude-oil-degrading thermo- philic bacterium isolated from an oil field. *Canadian Journal of Microbiology,* 14**,** 175–182.

HAQUE, M. M., GHOSH, T. S., KOMANDURI, D. & MANDE, S. S. 2009. SOrt-ITEMS: Sequence orthology based approach for improved taxonomic estimation of metagenomic sequences. *Bioinformatics,* 25**,** 1722-1730.

HARAYAMA, S. & TIMMIS, K. N. 2012. Catabolism ofAromatic Hydrocarbons by Pseudomonas. *Genetics of bacterial diversity***,** 151.

HASINGER, M., SCHERR, K. E., LUNDAA, T., BRÄUER, L., ZACH, C. & LOIBNER, A. P. 2012a. Changes in iso-and n-alkane distribution during biodegradation of crude oil under nitrate and sulphate reducing conditions. *Journal of biotechnology,* 157**,** 490-498.

HASINGER, M., SCHERR, K. E., LUNDAA, T., BRÄUER, L., ZACH, C. & LOIBNER, A. P. 2012b. Changes in iso-and< i> n</i>-alkane distribution during biodegradation of crude oil under nitrate and sulphate reducing conditions. *Journal of biotechnology,* 157**,** 490-498.

HEAD, I., MARTIN-JONES, D. & ROLING, W. 2006. Marine microorganisms make a meal of oil. *Nature Reviews in Microbiology,* 4**,** 173-182.

HEARN E M, PATEL D R, LEPORE B W, INDIC M. & B., A. 2009. Transmembrane passage of hydrophobic compounds through a protein channel wall. *Nature,* 458**,** 367–370.

HEBSGAARD, S. M., KORNING, P. G., TOLSTRUP, N., ENGELBRECHT, J., ROUZÉ, P. & BRUNAK, S. 1996. Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information. *Nucleic Acids Research,* 24**,** 3439-3452.

HEIDER, J. 2007. Adding handles to unhandy substrates: anaerobic hydrocarbon activation mechanisms. *Current opinion in chemical biology,* 11**,** 188-194.

HEIJTZ, R. D., WANG, S., ANUAR, F., QIAN, Y., BJÖRKHOLM, B., SAMUELSSON, A., HIBBERD, M. L., FORSSBERG, H. & PETTERSSON, S. 2011. Normal gut microbiota modulates brain development and behavior. *Proceedings of the National Academy of Sciences,* 108**,** 3047-3052.

HENSON, J., TISCHLER, G. & NING, Z. 2012. Next-generation sequencing and large genome assemblies. *Pharmacogenomics,* 13**,** 901-915.

HERMAN, P. L., BEHRENS, M., CHAKRABORTY, S., CHRASTIL, B. M., BARYCKI, J. & WEEKS, D. P. 2005. A Three-component Dicamba O-Demethylase from Pseudomonas maltophilia, Strain DI-6 GENE ISOLATION, CHARACTERIZATION, AND HETEROLOGOUS EXPRESSION. *Journal of Biological Chemistry,* 280**,** 24759-24767.

HERMUTH, K., LEUTHNER, B. & HEIDER, J. 2002. Operon structure and expression of the genes for benzylsuccinate synthase in Thauera aromatica strain K172. *Archives of microbiology,* 177**,** 132-138.

HERNÁNDEZ-ARRANZ, S., MORENO, R. & ROJO, F. 2013. The translational repressor Crc controls the Pseudomonas putida benzoate and alkane catabolic pathways using a multi-tier regulation strategy. *Environmental microbiology,* 15**,** 227-241.

HERWIG, R. P. 2011. Microorganisms that Degrade Oil and Bioremediation.

HIGASHI, S., BARRETO, A. M., CANTÃO, M. E. & DE VASCONCELOS, A. T. 2012. Analysis of composition-based metagenomic classification. *BMC genomics,* 13**,** S1.

HIRAISHI, A., KURAISHI, H. & KAWAHARA, K. 2000. Emendation of the description of Blastomonas natatoria (Sly 1985) Sly and Cahill 1997 as an aerobic photosynthetic bacterium and reclassification of Erythromonas ursincola Yurkov et al. 1997 as Blastomonas ursincola comb. nov. *International journal of systematic and evolutionary microbiology,* 50**,** 1113-1118.

HIRANO, S.-I., KITAUCHI, F., HARUKI, M., IMANAKA, T., MORIKAWA, M. & KANAYA, S. 2004. Isolation and characterization of Xanthobacter polyaromaticivorans sp. nov. 127W that degrades polycyclic and heterocyclic aromatic compounds under extremely low oxygen conditions. *Bioscience, biotechnology, and biochemistry,* 68**,** 557-564.

HOFF, K. J., LINGNER, T., MEINICKE, P. & TECH, M. 2009. Orphelia: predicting genes in metagenomic sequencing reads. *Nucleic acids research,* 37**,** W101-W105.

HOFF, K. J., TECH, M., LINGNER, T., DANIEL, R., MORGENSTERN, B. & MEINICKE, P. 2008. Gene prediction in metagenomic fragments: a large scale machine learning approach. *BMC bioinformatics,* 9**,** 217.

HOFFMAN, J., CLARK, M., AMOS, W. & PECK, L. 2012. Widespread amplification of amplified fragment length polymorphisms (AFLPs) in marine Antarctic animals. *Polar biology,* 35**,** 919-929.

HOOPER, L. V. & GORDON, J. I. 2001. Commensal host-bacterial relationships in the gut. *Science,* 292**,** 1115-1118.

HOU, C. T. 2006. 15 Biotransformation of Aliphatic Hydrocarbons and Fatty Acids.

HOUNSHELL, A. G. 2012. *Effect of biologically mediated sulfur oxidation on carbonate dissolution in submerged, fresh-water calcite caves.* University of Virginia.

HU, G., LI, J. & ZENG, G. 2013. Recent development in the treatment of oily sludge from petroleum industry: A review. *Journal of hazardous materials,* 261**,** 470-490.

HUA, F., WANG, H. Q., LI, Y. & ZHAO, Y. C. 2013. Trans-membrane transport of n-octadecane by Pseudomonas sp. DG17. *Journal of Microbiology,* 51**,** 791-799.

HUANG, T., HE, Z.-S., CUI, W.-R., CAI, Y.-D., SHI, X.-H., HU, L.-L. & CHOU, K.-C. 2013. A sequence-based approach for predicting protein disordered regions. *Protein and peptide letters,* 20**,** 243-248.

HUBÁLEK, T., VOSÁHLOVÁ, S., MATĚJŮ, V., KOVÁČOVÁ, N. & NOVOTNÝ, Č. 2007. Ecotoxicity monitoring of hydrocarbon-contaminated soil during bioremediation: a case study. *Archives of environmental contamination and toxicology,* 52**,** 1-7.

HUGENHOLTZ, P. & TYSON, G. 2008. Metagenomics. *Nature,* 455**,** 481-483.

HUNT, S., SMITH, K., HAMERTON, H. & SARGISSON, R. J. 2014. An Incident Control Centre in Action: Response to the Rena Oil Spill in New Zealand. *Journal of Contingencies and Crisis Management,* 22**,** 63-66.

HUSAIN, S. 2008. Microbial metabolism of high molecular weight polycyclic aromatic hydrocarbons. *Remediation,* 18**,** 131–161.

HUSI, H., SKIPWORTH, R. J., FEARON, K. C. & ROSS, J. A. 2013. LSCluster, a large-scale sequence clustering and aligning software for use in partial identity mapping and splice-variant analysis. *Journal of proteomics*.

HUSSAIN, T. & GONDAL, M. 2008. Monitoring and assessment of toxic metals in Gulf War oil spill contaminated soil using laser-induced breakdown spectroscopy. *Environmental monitoring and assessment,* 136**,** 391-399.

INSKEEP, W. P., RUSCH, D. B., JAY, Z. J., HERRGARD, M. J., KOZUBAL, M. A., RICHARDSON, T. H., MACUR, R. E., HAMAMURA, N., DEM JENNINGS, R. & FOUKE, B. W. 2010. Metagenomes from high-temperature chemotrophic systems reveal geochemical controls on microbial community structure and function. *PloS one,* 5**,** e9773.

ISHII Y, KONISHI J, OKADA H, HIRASAWA K, ONAKA T & M., S. 2000. Operon structure and functional analysis of the genes encoding thermophilic desulfurizing enzymes of Paenibacillus sp. A11-2. *Biochemical and Biophysical Research Communications,* 270**,** 81–82.

IWAI, S., CHAI, B., SUL, W. J., COLE, J. R., HASHSHAM, S. A. & TIEDJE, J. M. 2010. Gene-targeted-metagenomics reveals extensive diversity of aromatic dioxygenase genes in the environment. *The ISME journal,* 4**,** 279-285.

IWAI, S., JOHNSON, T. A., CHAI, B., HASHSHAM, S. A. & TIEDJE, J. M. 2011. Comparison of the specificities and efficacies of primers for aromatic dioxygenase gene analysis of environmental samples. *Applied and environmental microbiology,* 77**,** 3551-3557.

IWAKI, H., NAKAI, E., NAKAMURA, S. & HASEGAWA, Y. 2008. Isolation and Characterization of New Cyclohexylacetic Acid-Degrading Bacteria. *Current Microbiology,* 57**,** 107-110.

IZMALKOVA, T. Y., SAZONOVA, O. I., NAGORNIH, M. O., SOKOLOV, S. L., KOSHELEVA, I. A. & BORONIN, A. M. 2013a. The organization of naphthalene degradation genes in Pseudomonas putida strain AK5. *Research in microbiology,* 164**,** 244-253.

IZMALKOVA, T. Y., SAZONOVA, O. I., NAGORNIH, M. O., SOKOLOV, S. L., KOSHELEVA, I. A. & BORONIN, A. M. 2013b. The organization of naphthalene degradation genes in< i> Pseudomonas putida</i> strain AK5. *Research in microbiology,* 164**,** 244-253.

JABBOUR, D., SORGER, A., SAHM, K. & ANTRANIKIAN, G. 2013. A highly thermoactive and salt-tolerant α-amylase isolated from a pilot-plant biogas reactor. *Applied microbiology and biotechnology,* 97**,** 2971-2978.

JACQUOT, F., GUILIANO, M., DOUMENQ, P., MUNOZ, D. & MILLE, G. 1996. In vitro photooxidation of crude oil maltenic fractions: Evolution of fossil biomarkers and polycyclic aromatic hydrocarbons. *Chemosphere,* 33**,** 671-681.

JAYAMANI, I. & CUPPLES, A. M. 2015. Stable Isotope Probing and High-Throughput Sequencing Implicate Xanthomonadaceae and Rhodocyclaceae in Ethylbenzene Degradation. *Environmental Engineering Science*.

JAZESTANI, J. 2012. *Bioavailability and Biodegradation of Organic Xenobiotic Recalcitrant Polycyclic Aromatic Hydrocarbons (PAHs) in Different Soil Environments*, ProQuest, UMI Dissertations Publishing.

JI, Y., MAO, G., WANG, Y. & BARTLAM, M. 2013. Structural insights into diversity and n-alkane biodegradation mechanisms of alkane hydroxylases. *Frontiers in microbiology,* 4.

JIANG, H., CHEN, Y., JIANG, P., ZHANG, C., SMITH, T. J., MURRELL, J. C. & XING, X.-H. 2010. Methanotrophs: Multifunctional bacteria with promising applications in environmental bioengineering. *Biochemical Engineering Journal,* 49**,** 277-288.

JIANPING, X. 2006. Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. *Molecular Ecology,* 15**,** 1713–1731.

JIMÉNEZ, N., VIÑAS, M., BAYONA, J. M., ALBAIGES, J. & SOLANAS, A. M. 2007. The Prestige oil spill: bacterial community dynamics during a field biostimulation assay. *Applied microbiology and biotechnology,* 77**,** 935-945.

JIMESON, R. M., RADOSEVICH, M. C., STEVENS, R. R. & GIARDINO-RADOSEVICH, G. 2012. Biofuel petroleum fuel blend; gasoline and blend of C1-C8 alcohols. Google Patents.

JIN, H., CHEN, L., WANG, J. & ZHANG, W. 2014. Engineering biofuel tolerance in non-native producing microorganisms. *Biotechnology advances*.

JIN, H. M., KIM, J. M., LEE, H. J., MADSEN, E. L. & JEON, C. O. 2012. Alteromonas as a key agent of polycyclic aromatic hydrocarbon biodegradation in crude oil-contaminated coastal sediment. *Environmental science & technology,* 46**,** 7731-7740.

JIN, S., MAKRIS, T. M., BRYSON, T. A., SLIGAR, S. G. & DAWSON, J. H. 2003. Epoxidation of olefins by hydroperoxo-ferric cytochrome P450. *Journal of the American Chemical Society,* 125**,** 3406-3407.

JOBELIUS, C., RUTH, B., GRIEBLER, C., MECKENSTOCK, R. U., HOLLENDER, J., REINEKE, A., FRIMMEL, F. H. & ZWIENER, C. 2011. Metabolites indicate hot spots of biodegradation and biogeochemical gradients in a high-resolution monitoring well. *Environmental science & technology,* 45**,** 474.

JOBST, B., SCHÜHLE, K., LINNE, U. & HEIDER, J. 2010. ATP-dependent carboxylation of acetophenone by a novel type of carboxylase. *Journal of bacteriology*, 192**,** 1387-1394.

JOHNSON G. R., JAIN R. K. & C., S. J. 2002. Origins of the 2, 4-dinitrotoluene pathway. *Journal of bacteriology,* 184**,** 4219-4232.

JOHNSON, R. J., WEST, C. E., SWAIH, A. M., FOLWELL, B. D., SMITH, B. E., ROWLAND, S. J. & WHITBY, C. 2012. Aerobic biotransformation of alkyl branched aromatic alkanoic naphthenic acids via two different pathways by a new isolate of Mycobacterium. *Environmental microbiology,* 14**,** 872-882.

JONES, D. M., HEAD, I. M., GRAY, N. D., ADAMS, J. J., ROWAN, A. K., AITKEN, C. M., BENNETT, B., HUANG, H., BROWN, A., BOWLER, B. F. J., OLDENBURG, T. & ERDMANN, M. 2008. Crude-oil biodegradation via methanogenesis in subsurface petroleum reservoirs. *Nature,* 451**,** 176-180.

JONES, D. S., ALBRECHT, H. L., DAWSON, K. S., SCHAPERDOTH, I., FREEMAN, K. H., PI, Y., PEARSON, A. & MACALADY, J. L. 2011. Community genomic analysis of an extremely acidophilic sulfur-oxidizing biofilm. *The ISME journal,* 6**,** 158-170.

JOSHI, G., JUGRAN, J. & BHATT, J. 2014. Metagenomics: The Exploration of Unculturable Microbial World. *Advances in Biotechnology.* Springer.

JOUANNEAU, Y., MARTIN, F., KRIVOBOK, S. & WILLISON, J. C. 2011. Ring-hydroxylating dioxygenases involved in PAH biodegradation: structure, function and

biodiversity. *Microbial bioremediation of non metals: current research. Caister Academic Press, Norflok, UK***,** 149-175.

JULIAN, T. R. & SCHWAB, K. J. 2012. Challenges in environmental detection of human viral pathogens. *Current opinion in virology,* 2**,** 78-83.

JUTKINA, J., HANSEN, L. H., LI, L., HEINARU, E., VEDLER, E., JÕESAAR, M. & HEINARU, A. 2013. Complete nucleotide sequence of the self-transmissible TOL plasmid pD2RT provides new insight into arrangement of toluene catabolic plasmids. *Plasmid,* 70**,** 393-405.

KALYUZHNAYA, M., YANG, S., ROZOVA, O., SMALLEY, N., CLUBB, J., LAMB, A., GOWDA, G. N., RAFTERY, D., FU, Y. & BRINGEL, F. 2013. Highly efficient methane biocatalysis revealed in a methanotrophic bacterium. *Nature communications,* 4.

KANG, S., HERZBERG, M., RODRIGUES, D. F. & ELIMELECH, M. 2008. Antibacterial effects of carbon nanotubes: size does matter! *Langmuir,* 24**,** 6409-6413.

KANKAINEN, M., OJALA, T. & HOLM, L. 2012. BLANNOTATOR: enhanced homology-based function prediction of bacterial proteins. *BMC bioinformatics,* 13**,** 33.

KAPLAN, C. W. & KITTS, C. L. 2004. Bacterial succession in a petroleum land treatment unit. *Applied and Environmental Microbiology,* 70**,** 1777-1786.

KATARA, J., DESHMUKH, R., SINGH, N. K. & KAUR, S. 2013. Diversity Analysis of Bacillus thuringiensis Isolates Recovered from Diverse Habitats in India using Random Amplified Polymorphic DNA (RAPD) Markers. *Journal of Biological Sciences,* 13.

KATOH, K. & STANDLEY, D. M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution,* 30**,** 772-780.

KELLEY, D. R., LIU, B., DELCHER, A. L., POP, M. & SALZBERG, S. L. 2012. Gene prediction with Glimmer for metagenomic sequences augmented by classification and clustering. *Nucleic acids research,* 40**,** e9-e9.

KEYLOCK, C. 2005. Simpson diversity and the Shannon–Wiener index as special cases of a generalized entropy. *Oikos,* 109**,** 203-207.

KHAN, A., DE JONG, W., JANSENS, P. & SPLIETHOFF, H. 2009. Biomass combustion in fluidized bed boilers: potential problems and remedies. *Fuel processing technology,* 90**,** 21-50.

KHELIFI, N., GROSSI, V., HAMDI, M., DOLLA, A., THOLOZAN, J.-L., OLLIVIER, B. & HIRSCHLER-RÉA, A. 2010. Anaerobic oxidation of fatty acids and alkenes by the

hyperthermophilic sulfate-reducing archaeon Archaeoglobus fulgidus. *Applied and environmental microbiology,* 76**,** 3057-3060.

KIM, E.-Y., OH, K.-H., LEE, M.-H., KANG, C.-H., OH, T.-K. & YOON, J.-H. 2009. Novel cold-adapted alkaline lipase from an intertidal flat metagenome and proposal for a new family of bacterial lipases. *Applied and environmental microbiology,* 75**,** 257-260.

KIM, E. & ZYLSTRA, G. 1999. Functional analysis of genes involved in biphenyl, naphthalene, phenanthrene, and m-xylene degradation by Sphingomonas yanoikuyae B1. *Journal of Industrial Microbiology and Biotechnology,* 23**,** 294-302.

KIM, S.-J., KWEON, O., JONES, R. C., EDMONDSON, R. D. & CERNIGLIA, C. E. 2008. Genomic analysis of polycyclic aromatic hydrocarbon degradation in Mycobacterium vanbaalenii PYR-1. *Biodegradation,* 19**,** 859-881.

KIM, S.-J., KWEON, O., SUTHERLAND, J. B., KIM, H.-L., JONES, R. C., BURBACK, B. L., GRAVES, S. W., PSURNY, E. & CERNIGLIA, C. E. 2015. Dynamic response of Mycobacterium vanbaalenii PYR-1 to BP Deepwater Horizon crude oil. *Applied and environmental microbiology***,** AEM. 00730-15.

KING, R. W., BAUER, J. D. & BRADY, S. F. 2009. An Environmental DNA-Derived Type II Polyketide Biosynthetic Pathway Encodes the Biosynthesis of the Pentacyclic Polyketide Erdacin. *Angewandte Chemie International Edition,* 48**,** 6257-6261.

KIRCHER, M. & KELSO, J. 2010. High-throughput DNA sequencing–concepts and limitations. *Bioessays,* 32**,** 524-536.

KIRCHMAN, D. L. 2010. *Microbial ecology of the oceans*, John Wiley & Sons.

KIRIMURA K, HARADA K, IWASAWA H, TANAKA T, IWASAKI Y, FURUYA T, ISHII Y & K., K. 2004. Identification and functional analysis of the genes encoding dibenzothiophene-desulfurizing enzymes from thermophilic bacteria. *Applied microbial biotechnology,* 65**,** 703-13.

KLENK, H.-P. & GÖKER, M. 2010. En route to a genome-based classification of Archaea and Bacteria. *Systematic and applied microbiology,* 33**,** 175-182.

KLOER, D. P., HAGEL, C., HEIDER, J. & SCHULZ, G. E. 2006. Crystal Structure of Ethylbenzene Dehydrogenase from< i> Aromatoleum aromaticum</i>. *Structure,* 14**,** 1377-1388.

KNIEMEYER, O. & HEIDER, J. 2001a. Ethylbenzene Dehydrogenase, a Novel Hydrocarbon-oxidizing Molybdenum/Iron-Sulfur/Heme Enzyme. *Journal of Biological Chemistry,* 276**,** 21381-21386.

KNIEMEYER, O. & HEIDER, J. 2001b. (S)-1-Phenylethanol dehydrogenase of Azoarcus sp. strain EbN1, an enzyme of anaerobic ethylbenzene catabolism. *Archives of microbiology,* 176**,** 129-135.

KOBOLDT, D. C., STEINBERG, K. M., LARSON, D. E., WILSON, R. K. & MARDIS, E. R. 2013. The next-generation sequencing revolution and its impact on genomics. *Cell,* 155**,** 27-38.

KOCOT, K. M., CITARELLA, M. R., MOROZ, L. L. & HALANYCH, K. M. 2013. PhyloTreePruner: A Phylogenetic Tree-Based Approach for Selection of Orthologous Sequences for Phylogenomics. *Evolutionary bioinformatics online,* 9**,** 429.

KOELMANS, A., NOWACK, B. & WIESNER, M. 2009. Comparison of manufactured and black carbon nanoparticle concentrations in aquatic sediments. *Environmental pollution,* 157**,** 1110-1116.

KOMA, D., SAKASHITA, Y., KUBOTA, K., FUJII, Y., HASUMI, F., CHUNG, S. & KUBO, M. 2004. Degradation pathways of cyclic alkanes in Rhodococcus sp. NDKK48. *Applied microbiology and biotechnology,* 66**,** 92-99.

KONG, Y. 2011. Btrim: a fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. *Genomics,* 98**,** 152-153.

KÖSER, C. U., HOLDEN, M. T., ELLINGTON, M. J., CARTWRIGHT, E. J., BROWN, N. M., OGILVY-STUART, A. L., HSU, L. Y., CHEWAPREECHA, C., CROUCHER, N. J. & HARRIS, S. R. 2012. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *New England Journal of Medicine,* 366**,** 2267-2275.

KOSTICHKA, K., THOMAS, S. M., GIBSON, K. J., NAGARAJAN, V. & CHENG, Q. 2001a. Cloning and Characterization of a Gene Cluster for Cyclododecanone Oxidation in Rhodococcus ruber SC1. *The Journal of Bacteriology,* 183**,** 6478-6486.

KOSTICHKA, K., THOMAS, S. M., GIBSON, K. J., NAGARAJAN, V. & CHENG, Q. 2001b. Cloning and characterization of a gene cluster for cyclododecanone oxidation in Rhodococcus ruber SC1. *Journal of bacteriology,* 183**,** 6478-6486.

KOTANI, T., KAWASHIMA, Y., YURIMOTO, H., KATO, N. & SAKAI, Y. 2006. Gene structure and regulation of alkane monooxygenases in propane-utilizing Mycobacterium sp. TY-6 and Pseudonocardia sp. TY-7. *Journal of Bioscience and Bioengineering,* 102**,** 184-192.

KOTLAR, H. K., LEWIN, A., JOHANSEN, J., THRONE-HOLST, M., HAVERKAMP, T., MARKUSSEN, S., WINNBERG, A., RINGROSE, P., AAKVIK, T. & RYENG, E.

2011. High coverage sequencing of DNA from microorganisms living in an oil reservoir 2.5 kilometres subsurface. *Environmental Microbiology Reports,* 3**,** 674-681.

KOUKKOU, A.-I. & VANDERA, E. 2011. Hydrocarbon-degrading Soil Bacteria: Current Research. *Microbial Bioremediation of Non-metals: Current Research***,** 93.

KRAJEWSKI, S. S., JOSWIG, M., NAGEL, M. & NARBERHAUS, F. 2014. A tricistronic heat shock operon is important for stress tolerance of Pseudomonas putida and conserved in many environmental bacteria. *Environmental microbiology*.

KREISINGER, J. QIIME (Quantitative Insights Into Microbial Ecology): presenting of a software package for comparison and analysis of microbial communities, primarily based on high-throughput amplicon sequencing data. NGS workshop and conference MouseGene 2014, 2014.

KRELL, T., LACAL, J., GUAZZARONI, M. E., BUSCH, A., SILVA-JIMÉNEZ, H., FILLET, S., REYES-DARÍAS, J. A., MUÑOZ-MARTÍNEZ, F., RICO-JIMÉNEZ, M. & GARCÍA-FONTANA, C. 2012. Responses of Pseudomonas putida to toxic aromatic carbon sources. *Journal of biotechnology,* 160**,** 25-32.

KRISHNAKUMAR, A. M. 2007. *Structural studies of enzymes involved in propylene and acetone metabolism in Xanthobacter autotrophicus*, MONTANA STATE UNIVERSITY.

KROPP, K. G., DAVIDOVA, I. A. & SUFLITA, J. M. 2000. Anaerobic oxidation of n-dodecane by an addition reaction in a sulfate-reducing bacterial enrichment culture. *Applied and environmental microbiology,* 66**,** 5393-5398.

KU, C.-S. & ROUKOS, D. H. 2013. From next-generation sequencing to nanopore sequencing technology: paving the way to personalized genomic medicine. *Expert review of medical devices,* 10**,** 1-6.

KUBOTA, M., NODATE, M., YASUMOTO-HIROSE, M., UCHIYAMA, T., KAGAMI, O., SHIZURI, Y. & MISAWA, N. 2005. Isolation and functional analysis of cytochrome P450 CYP153A genes from various environments. *Bioscience, biotechnology, and biochemistry,* 69**,** 2421-2430.

KUROKI, T. 2012. Comparative Mutagenicity of Diol Epoxides of Benzo [a] pyrene and. *Molecular and cell biology,* 2**,** 123.

KWEON, O., KIM, S.-J., HOLLAND, R. D., CHEN, H., KIM, D.-W., GAO, Y., YU, L.-R., BAEK, S., BAEK, D.-H. & AHN, H. 2011. Polycyclic aromatic hydrocarbon-metabolic network in Mycobacterium vanbaalenii PYR-1. *Journal of bacteriology***,** JB. 00215-11.

LACAL, J., MUÑOZ-MARTÍNEZ, F., REYES-DARÍAS, J. A., DUQUE, E., MATILLA, M., SEGURA, A., CALVO, J. J. O., JÍMENEZ-SÁNCHEZ, C., KRELL, T. & RAMOS, J. L. 2011. Bacterial chemotaxis towards aromatic hydrocarbons in Pseudomonas. *Environmental microbiology,* 13**,** 1733-1744.

LADYGINA, N., DEDYUKHINA, E. & VAINSHTEIN, M. 2006. A review on microbial synthesis of hydrocarbons. *Process Biochemistry,* 41**,** 1001-1014.

LAI, Q., LI, W. & SHAO, Z. 2012. Complete genome sequence of Alcanivorax dieselolei type strain B5. *Journal of bacteriology,* 194**,** 6674-6674.

LANE, D. 1991. 16S/23S rRNA sequencing. *Nucleic acid techniques in bacterial systematics***,** 125-175.

LANG, S., SYLDATK, C. & RAU, U. 2000. 17 Enzymatic Synthesis and Modification of Glycolipids. *Enzymes in Lipid Modification,* 40**,** 363.

LARKIN, M. J., KULAKOV, L. A. & ALLEN, C. C. 2005. Biodegradation and Rhodococcus–masters of catabolic versatility. *Current Opinion in Biotechnology,* 16**,** 282-290.

LARTER, S. R. & HEAD, I. M. 2014. Oil sands and heavy oil: origin and exploitation. *Elements,* 10**,** 277-283.

LE BORGNEB, S. 2004. Petroleum biorefining: the selective removal of sulfur, nitrogen, and metals. *Petroleum Biotechnology: Developments and Perspectives,* 151**,** 29.

LEBLANC, J. G., MILANI, C., DE GIORI, G. S., SESMA, F., VAN SINDEREN, D. & VENTURA, M. 2013. Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Current opinion in biotechnology,* 24**,** 160-168.

LEE, E.-H. & CHO, K.-S. 2008. Characterization of cyclohexane and hexane degradation by Rhodococcus sp. EC1. *Chemosphere,* 71**,** 1738-1744.

LEE, K.-S., PARALES, J. V., FRIEMANN, R. & PARALES, R. E. 2005. Active site residues controlling substrate specificity in 2-nitrotoluene dioxygenase from Acidovorax sp. strain JS42. *Journal of Industrial Microbiology and Biotechnology,* 32**,** 465-473.

LEE, K., STOFFYN-EGLI, P., TREMBLAY, G. H., OWENS, E. H., SERGY, G. A., GUÉNETTE, C. C. & PRINCE, R. C. 2003. Oil–mineral aggregate formation on oiled beaches: natural attenuation and sediment relocation. *Spill Science & Technology Bulletin,* 8**,** 285-296.

LEES, J. G., LEE, D., STUDER, R. A., DAWSON, N. L., SILLITOE, I., DAS, S., YEATS, C., DESSAILLY, B. H., RENTZSCH, R. & ORENGO, C. A. 2014. Gene3D: Multi-domain annotations for protein sequence and comparative genome analysis. *Nucleic acids research,* 42**,** D240-D245.

LEGGETT, R. M., CLAVIJO, B. J., CLISSOLD, L., CLARK, M. D. & CACCAMO, M. 2013. NextClip: an analysis and read preparation tool for Nextera long mate pair libraries. *Bioinformatics***,** btt702.

LEHRBACH, P., WARD, J., MEULIEN, P. & BRODA, P. 1982. Physical mapping of TOL plasmids pWWO and pND2 and various R plasmid-TOL derivatives from Pseudomonas spp. *Journal of bacteriology,* 152**,** 1280-1283.

LEPAGE, P., LECLERC, M. C., JOOSSENS, M., MONDOT, S., BLOTTIÈRE, H. M., RAES, J., EHRLICH, D. & DORÉ, J. 2013. A metagenomic insight into our gut's microbiome. *Gut,* 62**,** 146-158.

LEUTHNER, B. & HEIDER, J. 2000a. Anaerobic toluene catabolism of Thauera aromatica: the bbs operon codes for enzymes of beta oxidation of the intermediate benzylsuccinate. *Journal of bacteriology,* 182**,** 272-277.

LEUTHNER, B. & HEIDER, J. 2000b. Anaerobic toluene catabolism of Thauera aromatica: the bbs operon codes for enzymes of β oxidation of the intermediate benzylsuccinate. *Journal of bacteriology,* 182**,** 272-277.

LEWIN, A., WENTZEL, A. & VALLA, S. 2012. Metagenomics of microbial life in extreme temperature environments. *Current Opinion in Biotechnology,* 24**,** 1–10.

LEWIN, A., WENTZEL, A. & VALLA, S. 2013. Metagenomics of microbial life in extreme temperature environments. *Current opinion in biotechnology,* 24**,** 516-525.

LEWIS, T. E., SILLITOE, I., ANDREEVA, A., BLUNDELL, T. L., BUCHAN, D. W., CHOTHIA, C., CUFF, A., DANA, J. M., FILIPPIS, I. & GOUGH, J. 2013. Genome3D: a UK collaborative project to annotate genomic sequences with predicted 3D structures based on SCOP and CATH domains. *Nucleic acids research,* 41**,** D499-D507.

LEY, R. E., PETERSON, D. A. & GORDON, J. I. 2006. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell,* 124**,** 837-848.

LEYS, N., RYNGAERT, A., BASTIAENS, L., TOP, E., VERSTRAETE, W. & SPRINGAEL, D. 2005. Culture independent detection of Sphingomonas sp. EPA 505 related strains in soils contaminated with polycyclic aromatic hydrocarbons (PAHs). *Microbial ecology,* 49**,** 443-450.

LI, L., LIU, X., YANG, W., XU, F., WANG, W., FENG, L., BARTLAM, M., WANG, L. & RAO, Z. 2008a. Crystal structure of long-chain alkane monooxygenase (LadA) in complex with coenzyme FMN: unveiling the long-chain alkane hydroxylase. *Journal of molecular biology,* 376**,** 453-465.

LI, L., LIU, X., YANG, W., XU, F., WANG, W., WANG, L., FENG, L., BARTLAM, M. & RAO, Z. 2008b. Crystal Structure of Long-Chain Alkane Monooxygenase (LadA) in Complex with Coenzyme FMN: Unveiling the Long-Chain Alkane Hydroxylase. *Journal of Molecular Biology*, 376**,** 453-465.

LI, R., ZHU, H., RUAN, J., QIAN, W., FANG, X., SHI, Z., LI, Y., LI, S., SHAN, G. & KRISTIANSEN, K. 2010. De novo assembly of human genomes with massively parallel short read sequencing. *Genome research,* 20**,** 265-272.

LI, S., ZHAO, H., LI, Y., NIU, S. & CAI, B. 2012. Complete genome sequence of the naphthalene-degrading Pseudomonas putida strain ND6. *Journal of bacteriology,* 194**,** 5154-5155.

LI, W., SHI, J., WANG, X., HAN, Y., TONG, W., MA, L., LIU, B. & CAI, B. 2004. Complete nucleotide sequence and organization of the naphthalene catabolic plasmid pND6-1 from Pseudomonas sp. strain ND6. *Gene,* 336**,** 231-240.

LI, Y., WANG, H., HUA, F., SU, M. & ZHAO, Y. 2014. Trans-membrane transport of fluoranthene by< i> Rhodococcus</i> sp. BAP-1 and optimization of uptake process. *Bioresource technology*.

LIANG, Q. & LLOYD-JONES, G. 2010. Sphingobium scionense sp. nov., an aromatic hydrocarbon-degrading bacterium isolated from contaminated sawmill soil. *International journal of systematic and evolutionary microbiology,* 60**,** 413-416.

LIANG, Y., VAN NOSTRAND, J. D., DENG, Y., HE, Z., WU, L., ZHANG, X., LI, G. & ZHOU, J. 2011. Functional gene diversity of soil microbial communities from five oil-contaminated fields in China. *The ISME journal,* 5**,** 403-413.

LIBRANDO, V. & PAPPALARDO, M. 2013. In silico bioremediation of polycyclic aromatic hydrocarbon: A frontier in environmental chemistry. *Journal of Molecular Graphics and Modelling,* 44**,** 1-8.

LILES, M. R., MANSKE, B. F., BINTRIM, S. B., HANDELSMAN, J. & GOODMAN, R. M. 2003. A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Applied and Environmental Microbiology,* 69**,** 2684-2691.

LIN, H., LIU, J.-Y., WANG, H.-B., AHMED, A. A. Q. & WU, Z.-L. 2011. Biocatalysis as an alternative for the production of chiral epoxides: a comparative review. *Journal of Molecular Catalysis B: Enzymatic,* 72**,** 77-89.

LIU, A., GARCIA-DOMINGUEZ, E., RHINE, E. D. & YOUNG, L. Y. 2004. A novel arsenate respiring isolate that can utilize aromatic substrates. *FEMS Microbiology Ecology,* 48**,** 323-332.

LIU, H., WANG, S.-J., ZHANG, J.-J., DAI, H., TANG, H. & ZHOU, N.-Y. 2011a. Patchwork assembly of nag-like nitroarene dioxygenase genes and the 3-chlorocatechol degradation cluster for evolution of the 2-chloronitrobenzene catabolism pathway in Pseudomonas stutzeri ZWLR2-1. *Applied and environmental microbiology,* 77**,** 4547-4552.

LIU, J.-F., SUN, X.-B., YANG, G.-C., MBADINGA, S. M., GU, J.-D. & MU, B.-Z. 2015. Analysis of microbial communities in the oil reservoir subjected to CO2-flooding by using functional genes as molecular biomarkers for microbial CO2 sequestration. *Frontiers in microbiology,* 6.

LIU, J., WANG, H., YANG, H., ZHANG, Y., WANG, J., ZHAO, F. & QI, J. 2013. Composition-based classification of short metagenomic sequences elucidates the landscapes of taxonomic and functional enrichment of microorganisms. *Nucleic acids research,* 41**,** e3-e3.

LIU, L., LI, Y., LI, S., HU, N., HE, Y., PONG, R., LIN, D., LU, L. & LAW, M. 2012. Comparison of next-generation sequencing systems. *BioMed Research International,* 2012.

LIU, T.-T., XU, Y., LIU, H., LUO, S., YIN, Y.-J., LIU, S.-J. & ZHOU, N.-Y. 2011b. Functional characterization of a gene cluster involved in gentisate catabolism in Rhodococcus sp. strain NCIMB 12038. *Applied microbiology and biotechnology,* 90**,** 671-678.

LIU, T.-T. & ZHOU, N.-Y. 2012. Novel L-cysteine-dependent maleylpyruvate isomerase in the gentisate pathway of Paenibacillus sp. strain NyZ101. *Journal of bacteriology,* 194**,** 3987-3994.

LIU, Y.-F., LIAO, R.-Z., DING, W.-J., YU, J.-G. & LIU, R.-Z. 2011c. Theoretical investigation of the first-shell mechanism of acetylene hydration catalyzed by a biomimetic tungsten complex. *Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry,* 16**,** 745-752.

LIU, Y., SCHMIDT, B. & MASKELL, D. L. 2011d. Parallelized short read assembly of large genomes using de Bruijn graphs. *BMC bioinformatics,* 12**,** 354.

LIU, Z., KLATT, C. G., WOOD, J. M., RUSCH, D. B., LUDWIG, M., WITTEKINDT, N., TOMSHO, L. P., SCHUSTER, S. C., WARD, D. M. & BRYANT, D. A. 2011e. Metatranscriptomic analyses of chlorophototrophs of a hot-spring microbial mat. *The ISME journal,* 5**,** 1279-1290.

LOGAN, A. C. & KATZMAN, M. 2005. Major depressive disorder: probiotics may be an adjuvant therapy. *Medical hypotheses,* 64**,** 533-538.

LOSS, R. A., FONTES, M. L., REGINATTO, V. & ANTÔNIO, R. V. 2013. Biohydrogen production by a mixed photoheterotrophic culture obtained from a Winogradsky column prepared from the sediment of a southern Brazilian lagoon. *Renewable Energy,* 50**,** 648-654.

LOZUPONE, C. & KNIGHT, R. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and environmental microbiology,* 71**,** 8228-8235.

LU, L., CHENG, S., GAO, J., GAO, G. & HE, M.-Y. 2007. Deep oxidative desulfurization of fuels catalyzed by ionic liquid in the presence of H2O2. *Energy & fuels,* 21**,** 383-384.

LUCIANI, F., BULL, R. A. & LLOYD, A. R. 2012. Next generation deep sequencing and vaccine design: today and tomorrow. *Trends in biotechnology,* 30**,** 443-452.

LUESKEN, F. A., VAN ALEN, T. A., VAN DER BIEZEN, E., FRIJTERS, C., TOONEN, G., KAMPMAN, C., HENDRICKX, T. L., ZEEMAN, G., TEMMINK, H. & STROUS, M. 2011. Diversity and enrichment of nitrite-dependent anaerobic methane oxidizing bacteria from wastewater sludge. *Applied microbiology and biotechnology,* 92**,** 845-854.

LUKASHIN, A. V. & BORODOVSKY, M. 1998. GeneMark. hmm: new solutions for gene finding. *Nucleic acids research,* 26**,** 1107-1115.

LUO, J. 2013. Applied Bioinformatics Tools. *Basics of Bioinformatics.* Springer.

MADANI, S. Y., MANDEL, A. & SEIFALIAN, A. M. 2013. A concise review of carbon nanotube's toxicology. *Nano reviews,* 4.

MAIL, E. Y., SANSCHAGRIN, S., BEAUMIER, D. & GREER, C. W. 2012. Metagenomic Analysis of the Bioremediation of Diesel- Contaminated Canadian High Arctic Soils. *PLoS ONE,* 7**,** 1-10.

MAITRA, R. D., KIM, J. & DUNBAR, W. B. 2012. Recent advances in nanopore sequencing. *Electrophoresis,* 33**,** 3418-3428.

MANI, D. & KUMAR, C. 2014. Biotechnological advances in bioremediation of heavy metals contaminated ecosystems: an overview with special reference to phytoremediation. *International Journal of Environmental Science and Technology,* 11**,** 843-872.

MANRAO, E. A., DERRINGTON, I. M., LASZLO, A. H., LANGFORD, K. W., HOPPER, M. K., GILLGREN, N., PAVLENOK, M., NIEDERWEIS, M. & GUNDLACH, J. H. 2012. Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nature biotechnology,* 30**,** 349-353.

MARGESIN, R. & SCHINNER, F. 2001a. Biodegradation and bioremediation of hydrocarbons in extreme environments. *Appl Microbiol Biotechnol,* 56**,** 650–663.

MARGESIN, R. & SCHINNER, F. 2001b. Biodegradation and bioremediation of hydrocarbons in extreme environments. *Applied microbiology and biotechnology,* 56**,** 650-663.

MARTIN, M. & RAHMANN, S. 2012. From cutadapt to sequencetools (sqt): a versatile toolset for sequencing projects. *EMBnet. journal,* 17**,** p. 35.

MARTINEZ, I., SANTOS, V. E., ALCON, A. & GARCIA-OCHOA, F. 2015. Enhancement of the biodesulfurization capacity of Pseudomonas putida CECT5279 by co-substrate addition. *Process Biochemistry,* 50**,** 119-124.

MASNADI, M. S., HABIBI, R., KOPYSCINSKI, J., HILL, J. M., BI, X., LIM, C. J., ELLIS, N. & GRACE, J. R. 2014. Fuel characterization and co-pyrolysis kinetics of biomass and fossil fuels. *Fuel,* 117**,** 1204-1214.

MASOUDI-NEJAD, A., NARIMANI, Z. & HOSSEINKHAN, N. 2013a. Emergence of Next-Generation Sequencing. *Next Generation Sequencing and Sequence Assembly.* Springer.

MASOUDI-NEJAD, A., NARIMANI, Z. & HOSSEINKHAN, N. 2013b. *Next Generation Sequencing and Sequence Assembly: Methodologies and Algorithms,* Springer.

MATSUBARA, T., OHSHIRO, T., NISHINA, Y. & IZUMI, Y. 2001. Purification, Characterization, and Overexpression of Flavin Reductase Involved in Dibenzothiophene Desulfurization byRhodococcus erythropolis D-1. *Applied and environmental microbiology,* 67**,** 1179-1184.

MATSUI, T., YAMAMOTO, T., SHINZATO, N., MITSUTA, T., NAKANO, K. & NAMIHIRA, T. 2014. Degradation of oil tank sludge using long-chain alkane-degrading bacteria. *Annals of Microbiology,* 64**,** 391-395.

MAUTER, M. S. & ELIMELECH, M. 2008. Environmental applications of carbon-based nanomaterials. *Environmental Science & Technology,* 42**,** 5843-5859.

MBADINGA, S. M., WANG, L.-Y., ZHOU, L., LIU, J.-F., GU, J.-D. & MU, B.-Z. 2011. Microbial communities involved in anaerobic degradation of alkanes. *International Biodeterioration & Biodegradation,* 65**,** 1-13.

MCCAMMICK, E., GOMASE, V., MCGENITY, T., TIMSON, D. & HALLSWORTH, J. 2010. Water-hydrophobic compound interactions with the microbial cell. *Handbook of Hydrocarbon and Lipid Microbiology.* Springer.

MCCARTHY, E. D. & CALVIN, K. 2008. Organic geochemical studies. I. Molecular criteria for hydrocarbon genesis.

MCGENITY, T. 2010. Halophilic Hydrocarbon Degraders. *Handbook of Hydrocarbon and Lipid Microbiology. Springer-Verlag Berlin Heidelberg***,** 1939-1951.

MECKENSTOCK, R. U. & MOUTTAKI, H. 2011. Anaerobic degradation of non-substituted aromatic hydrocarbons. *Current Opinion in Biotechnology,* 22**,** 406-414.

MEDINA, C. M. A., GARCÉS, A. A. A., ZULUAGA, I. C. G., ARDILA, M., ARIZA, G. & GUZMÁN, Z. 2014. Coexistence of Porphyromonas gingivalis, Tannerella forsythia and Treponema denticola in the Red Bacterial Complex in Chronic Periodontitis Subjects. *Int. J. Odontostomat,* 8**,** 359-364.

MEGHARAJ, M., RAMAKRISHNAN, B., VENKATESWARLU, K., SETHUNATHAN, N. & NAIDU, R. 2011. Bioremediation approaches for organic pollutants: a critical perspective. *Environment International,* 37**,** 1362-1375.

MERRIMAN, B., TORRENT, I., ROTHBERG, J. M. & TEAM, D. 2012. Progress in Ion Torrent semiconductor chip based sequencing. *Electrophoresis,* 33**,** 3397-3417.

METZKER, M. L. 2009. Sequencing technologies—the next generation. *Nature Reviews Genetics,* 11**,** 31-46.

MEYER, B. 2013. *Sulfur, energy, and environment*, Elsevier.

MICHAELIS, W., SEIFERT, R., NAUHAUS, K., TREUDE, T., THIEL, V., BLUMENBERG, M., KNITTEL, K., GIESEKE, A., PETERKNECHT, K., PAPE, T., BOETIUS, A., AMANN, R., JØRGENSEN, B. B., WIDDEL, F., PECKMANN, J., PIMENOV, N. V. & GULIN, M. B. 2002. Microbial reefs in the Black Sea fueled by anaerobic oxidation of methane. *Science (New York, N.Y.),* 297**,** 1013-1015.

MILAM, S. N., REYNOLDS, M. A., WELLINGTON, S. L. & BUHRMAN, F. A. 2014. Process for treating a hydrocarbon-containing feed. Google Patents.

MILANESI, L., D'ANGELO, D. & ROGOZIN, I. B. 1999. GeneBuilder: interactive in silico prediction of gene structure. *Bioinformatics,* 15**,** 612-621.

MILLER, R. & BARTHA, R. 1989. Evidence from tiposome encapsulation for transport-limited microbial metabolism of solid alkanes. *Applied Environmental Microbiology,* 55**,** 268-274.

MILLER, W. & MILLER, D. 1987. A micro-pipette method for soil mechanical analysis. *Communications in Soil Science & Plant Analysis.*

MINNECI, F., PIOVESAN, D., COZZETTO, D. & JONES, D. T. 2013. FFPred 2.0: Improved Homology-Independent Prediction of Gene Ontology Terms for Eukaryotic Protein Sequences. *PloS one,* 8**,** e63754.

MISHRA, S. K., SHRIVASTAV, A., MAURYA, R. R., PATIDAR, S. K., HALDAR, S. & MISHRA, S. 2012. Effect of light quality on the C-phycoerythrin production in marine cyanobacteria Pseudanabaena sp. isolated from Gujarat coast, India. *Protein expression and purification,* 81**,** 5-10.

MOCALI, S. & BENEDETTI, A. 2010. Exploring research frontiers in microbiology: the challenge of metagenomics in soil microbiology. *Research in microbiology,* 161**,** 497-505.

MOHAMMED, M. H., GHOSH, T. S., SINGH, N. K. & MANDE, S. S. 2011. SPHINX—an algorithm for taxonomic binning of metagenomic sequences. *Bioinformatics,* 27**,** 22-30.

MOHEBALI GASEMALI & ANDREW, B. S. 2008. Biocatalytic desulfurization (BDS) of petrodiesel fuels. *Microbioogy,* 154**,** 2169-2183.

MOKILI, J. L., ROHWER, F. & DUTILH, B. E. 2012. Metagenomics and future perspectives in virus discovery. *Current opinion in virology,* 2**,** 63-77.

MONSERUD, J. H. & SCHWARTZ, D. K. 2012. Effects of molecular size and surface hydrophobicity on oligonucleotide interfacial dynamics. *Biomacromolecules,* 13**,** 4002-4011.

MORAN, J. J., BEAL, E. J., VRENTAS, J. M., ORPHAN, V. J., FREEMAN, K. H. & HOUSE, C. H. 2008. Methyl sulfides as intermediates in the anaerobic oxidation of methane. *Environmental microbiology,* 10**,** 162-?

MORGAN, P. & WATKINSON, R. J. 1994. Biodegradation of components of petroleum. *Biochemistry of Microbial Degradation. Kluwer Academic Publishers, Dordrecht, Netherlands***,** 1-13.

MORIKAWA, M. 2010. Dioxygen activation responsible for oxidation of aliphatic and aromatic hydrocarbon compounds: current state and variants. *Applied microbiology and biotechnology,* 87**,** 1595-1603.

MOSHYNETS, O., BORETSKA, M. & SPIERS, A. J. 2013. From Winogradsky's column to contemporary research using bacterial microcosms. *Microcosms: ecology, biological implications and environmental impact*.

MOUGET, J.-L., ROSA, P. & TREMBLIN, G. 2004. Acclimation of Haslea ostrearia to light of different spectral qualities–confirmation ofchromatic adaptation'in diatoms. *Journal of Photochemistry and Photobiology B: Biology,* 75**,** 1-11.

MÜLLER, M. M., KÜGLER, J. H., HENKEL, M., GERLITZKI, M., HÖRMANN, B., PÖHNLEIN, M., SYLDATK, C. & HAUSMANN, R. 2012. Rhamnolipids—next generation surfactants? *Journal of biotechnology,* 162**,** 366-380.

MURATOVA, A. Y., TURKOVSKAYA, O., ANTONYUK, L., MAKAROV, O., POZDNYAKOVA, L. & IGNATOV, V. 2005. Oil-oxidizing potential of associative rhizobacteria of the genus Azospirillum. *Microbiology,* 74**,** 210-215.

MURRELL, J. & SMITH, T. 2010. Biochemistry and molecular biology of methane monooxygenase. *Handbook of Hydrocarbon and Lipid Microbiology***,** 1045-1055.

MUSAT, F., WILKES, H., BEHRENDS, A., WOEBKEN, D. & WIDDEL, F. 2010. Microbial nitrate-dependent cyclohexane degradation coupled with anaerobic ammonium oxidation. *The ISME journal,* 4**,** 1290-1301.

MYHR, S. & TORSVIK, T. 2000. Denitrovibrio acetiphilus, a novel genus and species of dissimilatory nitrate-reducing bacterium isolated from an oil reservoir model column. *International journal of systematic and evolutionary microbiology,* 50**,** 1611-1619.

NAKAMURA, S., MAEDA, N., MIRON, I. M., YOH, M., IZUTSU, K., KATAOKA, C., HONDA, T., YASUNAGA, T., NAKAYA, T. & KAWAI, J. 2008. Metagenomic diagnosis of bacterial infections. *Emerging infectious diseases,* 14**,** 1784.

NAKAMURA, S., YANG, C.-S., SAKON, N., UEDA, M., TOUGAN, T., YAMASHITA, A., GOTO, N., TAKAHASHI, K., YASUNAGA, T. & IKUTA, K. 2009. Direct metagenomic detection of viral pathogens in nasal and fecal specimens using an unbiased high-throughput sequencing approach. *PloS one,* 4**,** e4219.

NAKAR, D. & GUTNICK, D. L. 2001. Analysis of the wee gene cluster responsible for the biosynthesis of the polymeric bioemulsifier from the oil-degrading strain Acinetobacter lwoffii RAG-1. *Microbiology,* 147**,** 1937-1946.

NELSON, W. C., WOLLERMAN, L., BHAYA, D. & HEIDELBERG, J. F. 2011. Analysis of insertion sequences in thermophilic cyanobacteria: exploring the mechanisms of establishing, maintaining, and withstanding high insertion sequence abundance. *Applied and environmental microbiology,* 77**,** 5458-5466.

NERIA-GONZÁLEZ, I., WANG, E. T., RAMÍREZ, F., ROMERO, J. M. & HERNÁNDEZ-RODRÍGUEZ, C. 2006. Characterization of bacterial community associated to biofilms of corroded oil pipelines from the southeast of Mexico. *Anaerobe,* 12**,** 122-133.

NG, C., DEMAERE, M. Z., WILLIAMS, T. J., LAURO, F. M., RAFTERY, M., GIBSON, J. A., ANDREWS-PFANNKOCH, C., LEWIS, M., HOFFMAN, J. M. & THOMAS, T. 2010. Metaproteogenomic analysis of a dominant green sulfur bacterium from Ace Lake, Antarctica. *The ISME journal,* 4**,** 1002-1019.

NHI-CONG, L. T., MIKOLASCH, A., KLENK, H.-P. & SCHAUER, F. 2009. Degradation of the multiple branched alkane 2, 6, 10, 14-tetramethyl-pentadecane (pristane) in Rhodococcus ruber and Mycobacterium neoaurum. *International Biodeterioration & Biodegradation,* 63**,** 201-207.

NICOLOTTI, G. & EGLI, S. 1998. Soil contamination by crude oil: impact on the mycorrhizosphere and on the revegetation potential of forest trees. *Environmental Pollution,* 99**,** 37-43.

NIMCHUA, T., UENGWETWANIT, T. & EURWILAICHITR, L. 2012. Metagenomic analysis of novel lignocellulose-degrading enzymes from higher termite guts inhabiting microbes. *Journal of microbiology and biotechnology,* 22**,** 462-469.

NISHINO, S. F., SHIN, K. A., GOSSETT, J. M. & SPAIN, J. C. 2013. Cytochrome P450 Initiates Degradation of cis-Dichloroethene by Polaromonas sp. Strain JS666. *Applied and environmental microbiology,* 79**,** 2263-2272.

NISHINO, S. F. & SPAIN, J. C. 1995. Oxidative Pathway for the Biodegradation of Nitrobenzene by Comamonas sp. Strain JS765. *Applied and environmental microbiology,* 61**,** 2308-2313.

NOGUCHI, H., PARK, J. & TAKAGI, T. 2006. MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. *Nucleic acids research,* 34**,** 5623-5630.

NOLVAK, H., SILDVEE, T., KRIIPSALU, M. & TRUU, J. 2012. Application of microbial community profiling and functional gene detection for assessment of natural attenuation of petroleum hydrocarbons in boreal subsurface. *Boreal environment research,* 17**,** 113–127.

NORDIN, K., UNELL, M. & JANSSON, J. K. 2005. Novel 4-chlorophenol degradation gene cluster and degradation route via hydroxyquinol in Arthrobacter chlorophenolicus A6. *Applied and environmental microbiology,* 71**,** 6538-6544.

OBAYORI, S. & SALAM, L. B. 2010. Degradation of polycyclic aromatic hydrocarbons: role of plasmids. *Sci Res Essays,* 5**,** 4093-4106.

OBEROI, A. S., PHILIP, L. & BHALLAMUDI, S. M. 2015. Biodegradation of Various Aromatic Compounds by Enriched Bacterial Cultures: Part A–Monocyclic and

Polycyclic Aromatic Hydrocarbons. *Applied biochemistry and biotechnology,* 176**,** 1870-1888.

ODELL, P. R. 2013. *Oil and World Power (Routledge Revivals)*, Routledge.

ODOKUMA, L. & DICKSON, A. 2004. Bioremediation of a crude oil polluted tropical rain forest soil. *Global Journal of Environmental Sciences,* 2**,** 29-40.

OELSCHLÄGEL, M., GRÖNING, J. A., TISCHLER, D., KASCHABEK, S. R. & SCHLÖMANN, M. 2012. Styrene Oxide Isomerase of Rhodococcus opacus 1CP, a Highly Stable and Considerably Active Enzyme. *Applied and environmental microbiology,* 78**,** 4330-4337.

OHTSUBO, Y., NAGATA, Y., KIMBARA, K., TAKAGI, M. & OHTA, A. 2000. Expression of the bph genes involved in biphenyl/PCB degradation in Pseudomonas sp. KKS102 induced by the biphenyl degradation intermediate, 2-hydroxy-6-oxo-6-phenylhexa-2, 4-dienoic acid. *Gene,* 256**,** 223-228.

OLIVEIRA, C. R., FERREIRA, A. A., OLIVEIRA, C. J., AZEVEDO, D. A., SANTOS NETO, E. V. & AQUINO NETO, F. R. 2012. Biomarkers in crude oil revealed by comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry: Depositional paleoenvironment proxies. *Organic Geochemistry,* 46**,** 154-164.

OLIVEIRA, F. J. & DE FRANÇA, F. P. Increase in removal of polycyclic aromatic hydrocarbons during bioremediation of crude oil-contaminated sandy soil. Twenty-Sixth Symposium on Biotechnology for Fuels and Chemicals, 2005. Springer, 593-603.

ONO, A., MIYAZAKI, R. & SOTA, M. 2007. Isolation and characterization of naphthalene-catabolic genes and plasmids from oil-contaminated soil by using two cultivation-independent approaches. *Applied Microbiology and Biotechnology,* 74**,** 501-510.

ORCUTT, B. N., SYLVAN, J. B., KNAB, N. J. & EDWARDS, K. J. 2011. Microbial ecology of the dark ocean above, at, and below the seafloor. *Microbiology and Molecular Biology Reviews,* 75**,** 361-422.

ORTIZ DE MONTELLANO, P. R. 2010. Hydrocarbon hydroxylation by cytochrome P450 enzymes. *Chemical reviews,* 110**,** 932.

OSUJI, L. C. & ONOJAKE, C. M. 2004. Trace Heavy Metals Associated with Crude Oil: A Case Study of Ebocha-8 Oil-Spill-Polluted Site in Niger Delta, Nigeria. *Chemistry & biodiversity,* 1**,** 1708-1715.

PACWA-PŁOCINICZAK, M., PŁAZA, G. A., POLIWODA, A. & PIOTROWSKA-SEGET, Z. 2014. Characterization of hydrocarbon-degrading and biosurfactant-producing

Pseudomonas sp. P-1 strain as a potential tool for bioremediation of petroleum-contaminated soil. *Environmental Science and Pollution Research,* 21**,** 9385-9395.

PALFEY, B. A., BALLOU, D. P. & MASSEY, V. 1995. Oxygen activation by flavins and pterins. *Active Oxygen in Biochemistry.* Springer.

PANNARALE, P., CATALANO, D., DE CARO, G., GRILLO, G., LEO, P., PAPPADÀ, G., RUBINO, F., SCIOSCIA, G. & LICCIULLI, F. 2012. GIDL: a rule based expert system for GenBank Intelligent Data Loading into the Molecular Biodiversity database. *BMC bioinformatics,* 13**,** S4.

PAPIZADEH, M., ARDAKANI, M. R., MOTAMEDI, H., RASOULI, I. & ZAREI, M. 2011. C–S Targeted Biodegradation of Dibenzothiophene by Stenotrophomonas sp. NISOC-04. *Applied biochemistry and biotechnology,* 165**,** 938-948.

PARK, N., SHIRLEY, L., GU, Y., KEANE, T. M., SWERDLOW, H. & QUAIL, M. A. 2013. An improved approach to mate-paired library preparation for Illumina sequencing. *Methods in Next Generation Sequencing,* 1**,** 10-20.

PARK, S. R., HAN, A. R., BAN, Y.-H., YOO, Y. J., KIM, E. J. & YOON, Y. J. 2010. Genetic engineering of macrolide biosynthesis: past advances, current state, and future prospects. *Applied microbiology and biotechnology,* 85**,** 1227-1239.

PAUL, E. A. 2014. *Soil microbiology, ecology and biochemistry*, Academic press.

PAUZI ZAKARIA, M., OKUDA, T. & TAKADA, H. 2001. Polycyclic aromatic hydrocarbon (PAHs) and hopanes in stranded tar-balls on the coasts of Peninsular Malaysia: applications of biomarkers for identifying sources of oil pollution. *Marine Pollution Bulletin,* 42**,** 1357-1366.

PEARSON, W. R. 2014. BLAST and FASTA similarity searching for multiple sequence alignment. *Multiple Sequence Alignment Methods.* Springer.

PEREIRA, I. A. C., RAMOS, A. R., GREIN, F., MARQUES, M. C., DA SILVA, S. M. & VENCESLAU, S. S. 2011. A comparative genomic analysis of energy metabolism in sulfate reducing bacteria and archaea. *Frontiers in microbiology,* 2.

PERRY, L. L. & ZYLSTRA, G. J. 2007. Cloning of a gene cluster involved in the catabolism of p-nitrophenol by Arthrobacter sp. strain JS443 and characterization of the p-nitrophenol monooxygenase. *Journal of bacteriology,* 189**,** 7563-7572.

PERTEA, M., LIN, X. & SALZBERG, S. L. 2001. GeneSplicer: a new computational method for splice site prediction. *Nucleic acids research,* 29**,** 1185-1190.

PFENNIG, N. 1967. Photosynthetic bacteria. *Annual Reviews in Microbiology,* 21**,** 285-324.

PHALE, P. S., MAHAJAN, M. C. & VAIDYANATHAN, C. S. 2013. Biodegradation of polycyclic aromatic hydrocarbons. *Journal of the Indian Institute of Science,* 77**,** 141.

PICKRELL, W. O., REES, M. I. & CHUNG, S.-K. 2012. 1 Next Generation Sequencing Methodologies—An Overview. *Advances in Protein Chemistry and StructuralBiology,* 89**,** 1.

PIETROSKI, J. P., WHITE, J. R., DELAUNE, R. D., WANG, J. J. & DODLA, S. K. 2015. Fresh and weathered crude oil effects on potential denitrification rates of coastal marsh soil. *Chemosphere,* 134**,** 120-126.

PILLONI, G., VON NETZER, F., ENGEL, M. & LUEDERS, T. 2011. Electron acceptor-dependent identification of key anaerobic toluene degraders at a tar-oil-contaminated aquifer by Pyro-SIP. *FEMS Microbiology Ecology,* 78**,** 165-175.

POLZ, M. F., ALM, E. J. & HANAGE, W. P. 2013. Horizontal gene transfer and the evolution of bacterial and archaeal population structure. *Trends in Genetics,* 29**,** 170-175.

PRINCE, R., GRAMAIN, A. & MCGENITY, T. 2010. Prokaryotic hydrocarbon degraders. *Handbook of hydrocarbon and lipid microbiology.* Springer.

RABAUSCH, U., JUERGENSEN, J., ILMBERGER, N., BÖHNKE, S., FISCHER, S., SCHUBACH, B., SCHULTE, M. & STREIT, W. 2013. Functional Screening of Metagenome and Genome Libraries for Detection of Novel Flavonoid-Modifying Enzymes. *Applied and environmental microbiology,* 79**,** 4551-4563.

RABINOVITCH-DEERE, C. A. & PARALES, R. E. 2012. Three types of taxis used in the response of Acidovorax sp. strain JS42 to 2-nitrotoluene. *Applied and environmental microbiology,* 78**,** 2306-2315.

RABUS, R., KUBE, M., HEIDER, J., BECK, A., HEITMANN, K., WIDDEL, F. & REINHARDT, R. 2005. The genome sequence of an anaerobic aromatic-degrading denitrifying bacterium, strain EbN1. *Archives of microbiology,* 183**,** 27-36.

RABUS, R., WILKES, H., BEHRENDS, A., ARMSTROFF, A., FISCHER, T., PIERIK, A. J. & WIDDEL, F. 2001. Anaerobic initial reaction of n-alkanes in a denitrifying bacterium: evidence for (1-methylpentyl) succinate as initial product and for involvement of an organic radical in n-hexane metabolism. *Journal of bacteriology,* 183**,** 1707-1715.

RACHAMALLA, M. R., MONZOORUL, H. M. & MANDE, S. S. 2012. TWARIT: An extremely rapid and efficient approach for phylogenetic classification of metagenomic sequences. *Gene*.

RADFORD, A. D., CHAPMAN, D., DIXON, L., CHANTREY, J., DARBY, A. C. & HALL, N. 2012. Application of next-generation sequencing technologies in virology. *Journal of General Virology,* 93**,** 1853-1868.

RAGHUKUMAR, C., VIPPARTY, V., DAVID, J. & CHANDRAMOHAN, D. 2001. Degradation of crude oil by marine cyanobacteria. *Applied microbiology and biotechnology,* 57**,** 433-436.

RAHEB, J. 2011. The Study of Biodesulfurization Activity in Recombinant E. coli Strain by Cloning the dsz Genes Involve in 4S Pathway. *Journal of Sciences, Islamic Republic of Iran,* 22**,** 213-219.

RAHMAN, P., RAHMAN, T., LAKSHMANAPERUMALSAMY, P. & BANAT, I. M. 2002. Occurrence of crude oil degrading bacteria in gasoline and diesel station soils. *Journal of Basic Microbiology,* 42**,** 284-91.

RAJENDHRAN, J. & GUNASEKARAN, P. 2011. Microbial phylogeny and diversity: small subunit ribosomal RNA sequence analysis and beyond. *Microbiological Research,* 166**,** 99-110.

RAMADASS, K., SMITH, E., PALANISAMI, T., MATHIESON, G., SRIVASTAVA, P., MEGHARAJ, M. & NAIDU, R. 2015. Evaluation of constraints in bioremediation of weathered hydrocarbon-contaminated arid soils through microcosm biopile study. *International Journal of Environmental Science and Technology,* 12**,** 3597-3612.

RAMOS, R. T. J., CARNEIRO, A. R., CARACCIOLO, P. H., AZEVEDO, V., SCHNEIDER, M. P. C., BARH, D. & SILVA, A. 2013. Graphical contig analyzer for all sequencing platforms (G4ALL): a new stand-alone tool for finishing and draft generation of bacterial genomes. *Bioinformation,* 9**,** 599.

RANKIN, L. M., GOBSON, J., FRANZMANN, P. & BURTON, H. R. 1999. The chemical stratification and microbial communities of Ace Lake, Antarctica: A review of the characteristics of a marine-derived meromictic lake. *Polarforschung,* 66**,** 33-52.

RAO, T. A. A. 2011. *Designing a Barophile Enrichment Apparatus to Culture Deep Sea Microbes*, GRIN Verlag.

RASHEED, Z., RANGWALA, H. & BARBARA, D. LSH-Div: species diversity estimation using locality sensitive hashing. Bioinformatics and Biomedicine (BIBM), 2012 IEEE International Conference on, 2012. IEEE, 1-6.

RASHEED, Z., RANGWALA, H. & BARBARÁ, D. 2013. 16S rRNA metagenome clustering and diversity estimation using locality sensitive hashing. *BMC Systems Biology,* 7**,** 1-13.

RASKO, D. A., WEBSTER, D. R., SAHL, J. W., BASHIR, A., BOISEN, N., SCHEUTZ, F., PAXINOS, E. E., SEBRA, R., CHIN, C.-S. & ILIOPOULOS, D. 2011. Origins of the E. coli strain causing an outbreak of hemolytic–uremic syndrome in Germany. *New England Journal of Medicine,* 365**,** 709-717.

RATLEDGE, C. 1984. Microbial conversions of alkanes and fatty acids. *Journal of the American Oil Chemists' Society,* 61**,** 447-453.

RAYMENT, G. & HIGGINSON, F. R. 1992. *Australian laboratory handbook of soil and water chemical methods*, Inkata Press Pty Ltd.

REDMOND, M. C., VALENTINE, D. L. & SESSIONS, A. L. 2010. Identification of novel methane-, ethane-, and propane-oxidizing bacteria at marine hydrocarbon seeps by stable isotope probing. *Applied and environmental microbiology,* 76**,** 6412-6422.

REESE, M. G., EECKMAN, F. H., KULP, D. & HAUSSLER, D. 1997. Improved splice site detection in Genie. *Journal of computational biology,* 4**,** 311-323.

REINEKE, V., RULLKÖTTER, J., SMITH, E. L. & ROWLAND, S. J. 2006. Toxicity and compositional analysis of aromatic hydrocarbon fractions of two pairs of undegraded and biodegraded crude oils from the Santa Maria (California) and Vienna basins. *Organic geochemistry,* 37**,** 1885-1899.

RHEE, S.-K., CHANG, J. H., CHANG, Y. K. & CHANG, H. N. 1998. Desulfurization of dibenzothiophene and diesel oils by a newly isolated Gordona strain, CYKS1. *Applied and Environmental Microbiology,* 64**,** 2327-2331.

RIAZI, M. 2005. Characterization and Properties of Petroleum Fractions American Society for Testing and Materials (ASTM). *West Conshohocken, PA.*

RIOS-HERNANDEZ, L. A., GIEG, L. M. & SUFLITA, J. M. 2003. Biodegradation of an alicyclic hydrocarbon by a sulfate-reducing enrichment from a gas condensate-contaminated aquifer. *Applied and environmental microbiology,* 69**,** 434-443.

RISDON, G. C., POLLARD, S. J., BRASSINGTON, K. J., MCEWAN, J. N., PATON, G. I., SEMPLE, K. T. & COULON, F. D. R. 2008. Development of an analytical procedure for weathered hydrocarbon contaminated soils within a UK risk-based framework. *Analytical chemistry,* 80**,** 7090-7096.

RISER-ROBERTS, E. 1998. *Remediation of petroleum contaminated soils: biological, physical, and chemical processes*, CRC Press.

ROCHA, C. A., PEDREGOSA, A. M. & LABORDA, F. 2011. Biosurfactant-mediated biodegradation of straight and methyl-branched alkanes by Pseudomonas aeruginosa ATCC 55925. *AMB express,* 1**,** 1-10.

RODRÍGUEZ-HERVA, J. J., GARCÍA, V., HURTADO, A., SEGURA, A. & RAMOS, J. L. 2007. The ttgGHI solvent efflux pump operon of Pseudomonas putida DOT-T1E is located on a large self-transmissible plasmid. *Environmental microbiology,* 9**,** 1550-1561.

ROJO, F. 2009. Degradation of alkanes by bacteria. *Environmental microbiology,* 11**,** 2477-2490.

ROJO, F. 2010. Genetic Features and Regulation of n-Alkane Metabolism. *Handbook of Hydrocarbon and Lipid Microbiology.* Springer.

RONQUIST, F., TESLENKO, M., VAN DER MARK, P., AYRES, D. L., DARLING, A., HÖHNA, S., LARGET, B., LIU, L., SUCHARD, M. A. & HUELSENBECK, J. P. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology,* 61**,** 539-542.

RONTANI, J.-F., MOUZDAHIR, A., MICHOTEY, V. & BONIN, P. 2002. Aerobic and anaerobic metabolism of squalene by a denitrifying bacterium isolated from marine sediment. *Archives of microbiology,* 178**,** 279-287.

ROSEN, G. L., REICHENBERGER, E. R. & ROSENFELD, A. M. 2011. NBC: the Naive Bayes Classification tool webserver for taxonomic classification of metagenomic reads. *Bioinformatics,* 27**,** 127-129.

ROSENBERG, E. 2013. Hydrocarbon-Oxidizing Bacteria. *The Prokaryotes – Prokaryotic Physiology and Biochemistry***,** 201-214.

ROSENBERG, E., STACKEBRANDT, E., SCHLEIFER, K.-H., DWORKIN, M. & FALKOW, S. 2007. *Anaerobic Biodegradation of Hydrocarbons Including Methane,* New York, NY, Springer New York.

ROSNER, B. M. & SCHINK, B. 1995. Purification and characterization of acetylene hydratase of Pelobacter acetylenicus, a tungsten iron-sulfur protein. *Journal of bacteriology,* 177**,** 5767-5772.

ROSS, M. A. 1991. Environmental warfare and the Persian Gulf War: Possible remedies to combat intentional destruction of the environment. *Dick. J. Int'l L.,* 10**,** 515.

ROTARU, A. E., PROBIAN, C., WILKES, H. & HARDER, J. 2010. Highly enriched Betaproteobacteria growing anaerobically with p-xylene and nitrate. *FEMS microbiology ecology,* 71**,** 460-468.

ROUSK, J., BÅÅTH, E., BROOKES, P. C., LAUBER, C. L., LOZUPONE, C., CAPORASO, J. G., KNIGHT, R. & FIERER, N. 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. *The ISME journal,* 4**,** 1340-1351.

ROYO, J. L., HIDALGO, M. & RUIZ, A. 2007. Pyrosequencing protocol using a universal biotinylated primer for mutation detection and SNP genotyping. *Nature protocols,* 2**,** 1734-1739.

RUBY, J. G., BELLARE, P. & DERISI, J. L. 2013. PRICE: Software for the Targeted Assembly of Components of (Meta) Genomic Sequence Data. *G3: Genes*| *Genomes*| *Genetics,* 3**,** 865-880.

RUNDELL, E. A., BANTA, L. M., WARD, D. V., WATTS, C. D., BIRREN, B. & ESTEBAN, D. J. 2014. 16S rRNA Gene Survey of Microbial Communities in Winogradsky Columns. *PloS one,* 9**,** e104134.

SABEHI, G., MASSANA, R., BIELAWSKI, J. P., ROSENBERG, M., DELONG, E. F. & BÉJÀ, O. 2003. Novel proteorhodopsin variants from the Mediterranean and Red Seas. *Environmental microbiology,* 5**,** 842-849.

SAHARAN BS, SAHU RK & D, S. 2011. A Review on Biosurfactants: Fermentation, Current Developments and Perspectives. *Genetic Engineering and Biotechnology Journal,* 2011**,** 1-39.

SAIT, M., HUGENHOLTZ, P. & JANSSEN, P. H. 2002. Cultivation of globally distributed soil bacteria from phylogenetic lineages previously only detected in cultivation-independent surveys. *Environmental Microbiology,* 4**,** 654-666.

SAKAI, Y., MAENG, J. H., KUBOTA, S., TANI, A., TANI, Y. & KATO, N. 1996. A non-conventional dissimilation pathway for long chain n-alkanes in Acinetobacter sp. M-1 that starts with a dioxygenase reaction. *Journal of fermentation and bioengineering,* 81**,** 286-291.

SALAM, L., ILORI, M., AMUND, O., NUMATA, M., HORISAKI, T. & NOJIRI, H. 2014. Carbazole angular dioxygenation and mineralization by bacteria isolated from hydrocarbon-contaminated tropical African soil. *Environmental Science and Pollution Research,* 21**,** 9311-9324.

SÁNCHEZ-PALENCIA GONZÁLEZ, Y. 2011. Bioremediation of oil spills.

SATHISHKUMAR, M. B., ARTHUR RAJ, BAIK, S.-H. & YUN, S.-E. 2008. Biodegradation of Crude Oil by Individual Bacterial Strains and a Mixed Bacterial Consortium Isolated from Hydrocarbon Contaminated Areas. *Clean,* 36**,** 92 – 96.

SATHYA, T., JACOB, A. M. & KHAN, M. 2014. Cloning and molecular modelling of pectin degrading glycosyl hydrolase of family 28 from soil metagenomic library. *Molecular biology reports***,** 1-12.

SAYARA, T., BORRÀS, E., CAMINAL, G., SARRÀ, M. & SÁNCHEZ, A. 2011. Bioremediation of PAHs-contaminated soil through composting: Influence of bioaugmentation and biostimulation on contaminant biodegradation. *International Biodeterioration & Biodegradation,* 65**,** 859-865.

SAYAVEDRA-SOTO, L., HAMAMURA, N., LIU, C., KIMBREL, J. & CHANG JH, A. D. 2011. The membrane-associated monooxygenase in the butane-oxidizing Gram-positive bacterium Nocardioides sp. strain CF8 is a novel member of the AMO/PMO family. *Environmental Microbiology Reports,* 3**,** 390–396.

SAYAVEDRA-SOTO, L. A., HAMAMURA, N., LIU, C. W., KIMBREL, J. A., CHANG, J. H. & ARP, D. J. 2011. The membrane-associated monooxygenase in the butane-oxidizing Gram-positive bacterium Nocardioides sp. strain CF8 is a novel member of the AMO/PMO family. *Environmental microbiology reports,* 3**,** 390-396.

SCHELLER, S., GOENRICH, M., BOECHER, R., THAUER, R. K. & JAUN, B. 2010. The key nickel enzyme of methanogenesis catalyses the anaerobic oxidation of methane. *Nature,* 465**,** 606-608.

SCHIMEL, J., BALSER, T. C. & WALLENSTEIN, M. 2007. Microbial stress-response physiology and its implications for ecosystem function. *Ecology,* 88**,** 1386-1394.

SCHINK, B. 1985. Degradation @@of unsaturated hydrocarbons by methanogenic enrichment cultures. *FEMS Microbiology Letters,* 31**,** 69-77.

SCHINK, B., JANSSEN, P. & BRUNE, A. 2011. Bergey's Manual of Systematic Bacteriology. Springer.

SCHIRMER, A., RUDE, M. A., LI, X., POPOVA, E. & DEL CARDAYRE, S. B. 2010. Microbial biosynthesis of alkanes. *Science,* 329**,** 559-562.

SCHLOSS, P. D. & HANDELSMAN, J. 2003. Biotechnological prospects from metagenomics. *Current Opinion in Biotechnology,* 14**,** 303-10.

SCHLOSS, P. D. & HANDELSMAN, J. 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Applied and environmental microbiology,* 71**,** 1501-1506.

SCHLOSS, P. D., WESTCOTT, S. L., RYABIN, T., HALL, J. R., HARTMANN, M., HOLLISTER, E. B., LESNIEWSKI, R. A., OAKLEY, B. B., PARKS, D. H. &

ROBINSON, C. J. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology,* 75**,** 7537-7541.

SCHMALENBERGER, A. & TEBBE, C. C. 2014. Profiling the Diversity of Microbial Communities with Single-Strand Conformation Polymorphism (SSCP). *Environmental Microbiology.* Springer.

SCHNEIKER, S., DOS SANTOS, V. A. M., BARTELS, D., BEKEL, T., BRECHT, M., BUHRMESTER, J., CHERNIKOVA, T. N., DENARO, R., FERRER, M. & GERTLER, C. 2006. Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium Alcanivorax borkumensis. *Nature biotechnology,* 24**,** 997-1004.

SCHULER, L., JOUANNEAU, Y., CHADHAIN, S. M. N., MEYER, C., POULI, M., ZYLSTRA, G. J., HOLS, P. & AGATHOS, S. N. 2009. Characterization of a ring-hydroxylating dioxygenase from phenanthrene-degrading Sphingomonas sp. strain LH128 able to oxidize benz [a] anthracene. *Applied microbiology and biotechnology,* 83**,** 465-475.

SCHURIG, C., MILTNER, A. & KAESTNER, M. 2014. Hexadecane and pristane degradation potential at the level of the aquifer—evidence from sediment incubations compared to in situ microcosms. *Environmental Science and Pollution Research***,** 1-14.

SCHWARTZ, S., ZHANG, Z., FRAZER, K. A., SMIT, A., RIEMER, C., BOUCK, J., GIBBS, R., HARDISON, R. & MILLER, W. 2000. PipMaker—a web server for aligning two genomic DNA sequences. *Genome Research,* 10**,** 577-586.

SCOTT-FORDSMAND, J. J., KROGH, P. H., SCHAEFER, M. & JOHANSEN, A. 2008. The toxicity testing of double-walled nanotubes-contaminated food to< i> Eisenia veneta</i> earthworms. *Ecotoxicology and environmental safety,* 71**,** 616-619.

SEELIGER, S., CORD-RUWISCH, R. & SCHINK, B. 1998. A periplasmic and extracellular c-type cytochrome of Geobacter sulfurreducens acts as a ferric iron reductase and as an electron carrier to other acceptors or to partner bacteria. *Journal of bacteriology,* 180**,** 3686-3691.

SEGURA, A., MOLINA, L., FILLET, S., KRELL, T., BERNAL, P., MUÑOZ-ROJAS, J. & RAMOS, J.-L. 2012. Solvent tolerance in Gram-negative bacteria. *Current opinion in biotechnology,* 23**,** 415-421.

SEIFERT, J., TAUBERT, M., JEHMLICH, N., SCHMIDT, F., VÖLKER, U., VOGT, C., RICHNOW, H. H. & VON BERGEN, M. 2012. Protein-based stable isotope probing (protein-SIP) in functional metaproteomics. *Mass spectrometry reviews,* 31**,** 683-697.

SEITZ, E. M., BROCKMAN, J. P., SANDLER, S. J., CLARK, A. J. & KOWALCZYKOWSKI, S. C. 1998. RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange. *Genes & development,* 12**,** 1248-1253.

SELESI, D., JEHMLICH, N., VON BERGEN, M., SCHMIDT, F., RATTEI, T., TISCHLER, P., LUEDERS, T. & MECKENSTOCK, R. U. 2009. Combined genomic and proteomic approaches identify gene clusters involved in anaerobic 2-methylnaphthalene degradation in the sulfate-reducing enrichment culture N47. *Journal of bacteriology,* 192**,** 295-306.

SELESI, D., JEHMLICH, N., VON BERGEN, M., SCHMIDT, F., RATTEI, T., TISCHLER, P., LUEDERS, T. & MECKENSTOCK, R. U. 2010. Combined genomic and proteomic approaches identify gene clusters involved in anaerobic 2-methylnaphthalene degradation in the sulfate-reducing enrichment culture N47. *Journal of bacteriology,* 192**,** 295-306.

SELESI, D. & MECKENSTOCK, R. U. 2009a. Anaerobic degradation of the aromatic hydrocarbon biphenyl by a sulfate-reducing enrichment culture. *FEMS microbiology ecology,* 68**,** 86-93.

SELESI, D. E. & MECKENSTOCK, R. U. 2009b. Anaerobic degradation of the aromatic hydrocarbon biphenyl by a sulfate-reducing enrichment culture. *FEMS Microbiology Ecology,* 68**,** 86-93.

SEO, J.-S. 2012. Multiple pathways in the degradation of dibenzothiophene by Mycobacterium aromativorans strain JS19b1T. *Journal of the Korean Society for Applied Biological Chemistry,* 55**,** 613-618.

SEO, J.-S., KEUM, Y.-S., HU, Y., LEE, S.-E. & LI, Q. X. 2006. Phenanthrene degradation in Arthrobacter sp. P1-1: Initial 1, 2-, 3, 4-and 9, 10-dioxygenation, and meta-and ortho-cleavages of naphthalene-1, 2-diol after its formation from naphthalene-1, 2-dicarboxylic acid and hydroxyl naphthoic acids. *Chemosphere,* 65**,** 2388-2394.

SEO, J., KANG, S.-I., RYU, J.-Y., LEE, Y.-J., PARK, K. D., KIM, M., WON, D., PARK, H.-Y., AHN, J.-H. & CHONG, Y. 2010. Location of flavone B-ring controls regioselectivity and stereoselectivity of naphthalene dioxygenase from Pseudomonas sp. strain NCIB 9816-4. *Applied microbiology and biotechnology,* 86**,** 1451-1462.

SHAH, V., ZAKRZEWSKI, M., WIBBERG, D., EIKMEYER, F., SCHLÜTER, A. & MADAMWAR, D. 2013. Taxonomic Profiling and Metagenome Analysis of a Microbial Community from a Habitat Contaminated with Industrial Discharges. *Microbial ecology,* 66**,** 533-550.

SHAPIRO, J., CHARBIT, A., BENSON, S., CARUSO, M., LAUX, R., MEYER, R. & BANUETT, F. 1981. Perspectives for genetic engineering of hydrocarbon oxidizing bacteria. *Trends in the Biology of Fermentations for Fuels and Chemicals.* Springer.

SHARMA, S., KHAN, F. G. & QAZI, G. N. 2010. Molecular cloning and characterization of amylase from soil metagenomic library derived from Northwestern Himalayas. *Applied microbiology and biotechnology,* 86**,** 1821-1828.

SHAVANDI, M., SOHEILI, M., ZAREIAN, S., AKBARI, N. & KHAJEH, K. 2013. THE GENE CLONING, OVEREXPRESSION, PURIFICATION, AND CHARACTERIZATION OF DIBENZOTHIOPHENE MONOOXYGENASE AND DESULFINASE FROM GORDONIA ALKANIVORANS RIPI90A. *Journal of Petroleum Science and Technology,* 3**,** 57-64.

SHE, R., CHU, J. S.-C., WANG, K., PEI, J. & CHEN, N. 2009. GenBlastA: enabling BLAST to identify homologous gene sequences. *Genome research,* 19**,** 143-149.

SHENDURE, J. & JI, H. 2008. Next-generation DNA sequencing. *Nature biotechnology,* 26**,** 1135-1145.

SHENDURE, J. A., PORRECA, G. J., CHURCH, G. M., GARDNER, A. F., HENDRICKSON, C. L., KIELECZAWA, J. & SLATKO, B. E. 2011. Overview of DNA sequencing strategies. *Current Protocols in Molecular Biology***,** 7.1. 1-7.1. 23.

SHENNAN, J. L. 2006. Utilisation of C2–C4 gaseous hydrocarbons and isoprene by microorganisms. *Journal of Chemical Technology and Biotechnology,* 81**,** 237-256.

SHERRY, A., GRAY, N., DITCHFIELD, A., AITKEN, C., JONES, D., RÖLING, W., HALLMANN, C., LARTER, S., BOWLER, B. & HEAD, I. 2013. Anaerobic biodegradation of crude oil under sulphate-reducing conditions leads to only modest enrichment of recognized sulphate-reducing taxa. *International Biodeterioration & Biodegradation,* 81**,** 105-113.

SHI, Q., HOU, D., CHUNG, K. H., XU, C., ZHAO, S. & ZHANG, Y. 2010. Characterization of heteroatom compounds in a crude oil and its saturates, aromatics, resins, and asphaltenes (SARA) and non-basic nitrogen fractions analyzed by negative-ion electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Energy & Fuels,* 24**,** 2545-2553.

SHI, S., QU, Y., MA, F. & ZHOU, J. 2014. Bioremediation of coking wastewater containing carbazole, dibenzofuran, dibenzothiophene and naphthalene by a naphthalene-cultivated Arthrobacter sp. W1. *Bioresource technology,* 164**,** 28-33.

SHIMOYAMA, T., YAMAZAWA, A., UENO, Y. & WATANABE, K. 2009. Short Communication Phylogenetic Analyses of Bacterial Communities Developed in a Cassette-Electrode Microbial Fuel Cell. *Microbes Environ,* 24**,** 188-192.

SHINDO, K., NAKAMURA, R., OSAWA, A., KAGAMI, O., KANOH, K., FURUKAWA, K. & MISAWA, N. 2005. Biocatalytic synthesis of monocyclic arene-dihydrodiols and -diols by Escherichia coli cells expressing hybrid toluene/biphenyl dioxygenase and dihydrodiol dehydrogenase genes. *Journal of Molecular Catalysis B: Enzymatic,* 35**,** 134-141.

SHINODA, Y., SAKAI, Y., UENISHI, H., UCHIHASHI, Y., HIRAISHI, A., YUKAWA, H., YURIMOTO, H. & KATO, N. 2004. Aerobic and Anaerobic Toluene Degradation by a Newly Isolated Denitrifying Bacterium, Thauera sp. Strain DNT-1. *Applied and Environmental Microbiology,* 70**,** 1385-1392.

SHIRANI, M., MIRVAGHEFI, A., FARAHMAND, H. & ABDOLLAHI, M. 2012. Biomarker responses in mudskipper (Periophthalmus waltoni) from the coastal areas of the Persian Gulf with oil pollution. *Environmental toxicology and pharmacology,* 34**,** 705-713.

SHOKRALLA, S., SPALL, J. L., GIBSON, J. F. & HAJIBABAEI, M. 2012. Next-generation sequencing technologies for environmental DNA research. *Molecular Ecology,* 21**,** 1794-1805.

SIEGERT, M., SITTE, J., GALUSHKO, A. & KRÜGER, M. 2014. Starting up microbial enhanced oil recovery. *Geobiotechnology II.* Springer.

SIERRA-GARCÍA, I. N., ALVAREZ, J. C., DE VASCONCELLOS, S. P., DE SOUZA, A. P., DOS SANTOS NETO, E. V. & DE OLIVEIRA, V. M. 2014. New Hydrocarbon Degradation Pathways in the Microbial Metagenome from Brazilian Petroleum Reservoirs. *PloS one,* 9**,** e90087.

SIKKEMA, J., DE BONT, J. & POOLMAN, B. 1995. Mechanisms of membrane toxicity of hydrocarbons. *Microbiological reviews,* 59**,** 201-222.

SILVA, T., VERDE, L., NETO, E. S. & OLIVEIRA, V. 2013. Diversity analyses of microbial communities in petroleum samples from Brazilian oil fields. *International Biodeterioration & Biodegradation,* 81**,** 57-70.

SIMON, C. & DANIEL, R. 2011. Metagenomic analyses: past and future trends. *Applied and environmental microbiology,* 77**,** 1153-1161.

SINGH, A., VAN HAMME, J. D., KUHAD, R. C., PARMAR, N. & WARD, O. P. 2014a. Subsurface Petroleum Microbiology. *Geomicrobiology and Biogeochemistry.* Springer.

SINGH, D., KUMARI, A. & RAMANATHAN, G. 2014b. 3-Nitrotoluene dioxygenase from Diaphorobacter sp. strains: cloning, sequencing and evolutionary studies. *Biodegradation,* 25**,** 479-492.

SINGH, J., BEHAL, A., SINGLA, N., JOSHI, A., BIRBIAN, N., SINGH, S., BALI, V. & BATRA, N. 2009. Metagenomics: Concept, methodology, ecological inference and recent advances. *Biotechnology Journal,* 4**,** 480-94.

SJÖBERG, F., NOWROUZIAN, F., RANGEL, I., HANNOUN, C., MOORE, E., ADLERBERTH, I. & WOLD, A. E. 2013. Comparison between terminal-restriction fragment length polymorphism (T-RFLP) and quantitative culture for analysis of infants' gut microbiota. *Journal of microbiological methods,* 94**,** 37-46.

SLOBODKINA, G., KOLGANOVA, T., CHERNYH, N., QUERELLOU, J., BONCH-OSMOLOVSKAYA, E. & SLOBODKIN, A. 2009. Deferribacter autotrophicus sp. nov., an iron (III)-reducing bacterium from a deep-sea hydrothermal vent. *International journal of systematic and evolutionary microbiology,* 59**,** 1508-1512.

SMALL, F. J. & ENSIGN, S. A. 1995. Carbon dioxide fixation in the metabolism of propylene and propylene oxide by Xanthobacter strain Py2. *Journal of bacteriology,* 177**,** 6170-6175.

SMEDS, L. & KÜNSTNER, A. 2011. ConDeTri-a content dependent read trimmer for Illumina data. *PloS one,* 6**,** e26314.

SMITH, S. M., RAWAT, S., TELSER, J., HOFFMAN, B. M., STEMMLER, T. L. & ROSENZWEIG, A. C. 2011. Crystal structure and characterization of particulate methane monooxygenase from Methylocystis species strain M. *Biochemistry,* 50**,** 10231-10240.

SMITH, T. J. 2010. Monooxygenases, Bacterial: Oxidation of Alkenes. *Encyclopedia of Industrial Biotechnology*.

SMITS, T. H., BALADA, S. B., WITHOLT, B. & VAN BEILEN, J. B. 2002. Functional analysis of alkane hydroxylases from gram-negative and gram-positive bacteria. *Journal of bacteriology,* 184**,** 1733-1742.

SOBERÓN-CHÁVEZ, G. & MAIER, R. M. 2011. Biosurfactants: a general overview. *Biosurfactants.* Springer.

SOERGEL, D. A., DEY, N., KNIGHT, R. & BRENNER, S. E. 2012. Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. *The ISME journal,* 6**,** 1440-1444.

SOLEIMANI, M., BASSI, A. & MARGARITIS, A. 2007. Biodesulfurization of refractory organic sulfur compounds in fossil fuels. *Biotechnology advances,* 25**,** 570-596.

SOLOMON, K. V., HAITJEMA, C. H., THOMPSON, D. A. & O'MALLEY, M. A. 2014. Extracting data from the muck: deriving biological insight from complex microbial communities and non-model organisms with next generation sequencing. *Current Opinion in Biotechnology,* 28**,** 103-110.

SONOMOTO, K., HOQ, M. M., TANAKA, A. & FUKUI, S. 1983. 11β-Hydroxylation of cortexolone (Reichstein compound S) to hydrocortisone by Curvularia lunata entrapped in photo-cross-linked resin gels. *Applied and environmental microbiology,* 45**,** 436-443.

SPAIN, J. C. & GIBSON, D. T. 1991. Pathway for biodegradation of p-nitrophenol in a Moraxella sp. *Applied and Environmental Microbiology,* 57**,** 812-819.

SPANGGORD, R. J., SPAIN, J., NISHINO, S. & MORTELMANS, K. 1991. Biodegradation of 2, 4-dinitrotoluene by a Pseudomonas sp. *Applied and Environmental Microbiology,* 57**,** 3200-3205.

SPARKS, D. L. 2003. *Environmental soil chemistry,* Academic press.

SPROCATI, A. R., ALISI, C., TASSO, F., MARCONI, P., SCIULLO, A., PINTO, V., CHIAVARINI, S., UBALDI, C. & CREMISINI, C. 2012. Effectiveness of a microbial formula, as a bioaugmentation agent, tailored for bioremediation of diesel oil and heavy metal co-contaminated soil. *Process Biochemistry,* 47**,** 1649-1655.

STACHLER, E. & BIBBY, K. 2014. Metagenomic Evaluation of the Highly Abundant Human Gut Bacteriophage CrAssphage for Source Tracking of Human Fecal Pollution. *Environmental Science & Technology Letters,* 1**,** 405-409.

STEELE, H. L. & STREIT, W. R. 2005. Metagenomics: advances in ecology and biotechnology. *FEMS Microbiology Letters,* 247**,** 105-111.

STEFANIS, C., ALEXOPOULOS, A., VOIDAROU, C., VAVIAS, S. & BEZIRTZOGLOU, E. 2013. Principal methods for isolation and identification of soil microbial communities. *Folia microbiologica,* 58**,** 61-68.

STRANNEHEIM, H. & LUNDEBERG, J. 2012. Stepping stones in DNA sequencing. *Biotechnology Journal,* 7**,** 1063-1073.

STURCHIO, N. C., BÖHLKE, J. K., BELOSO, A. D., STREGER, S. H., HERATY, L. J. & HATZINGER, P. B. 2007. Oxygen and chlorine isotopic fractionation during perchlorate biodegradation: Laboratory results and implications for forensics and natural attenuation studies. *Environmental science & technology,* 41**,** 2796-2802.

SU, C., LEI, L., DUAN, Y., ZHANG, K.-Q. & YANG, J. 2012. Culture-independent methods for studying environmental microorganisms: methods, application, and perspective. *Applied microbiology and biotechnology,* 93**,** 993-1003.

SUBRAMANIAN, A. R., KAUFMANN, M. & MORGENSTERN, B. 2008. DIALIGN-TX: greedy and progressive approaches for segment-based multiple sequence alignment. *Algorithms Mol Biol,* 3.

SUJA, F., RAHIM, F., TAHA, M. R., HAMBALI, N., RIZAL RAZALI, M., KHALID, A. & HAMZAH, A. 2014. Effects of local microbial bioaugmentation and biostimulation on the bioremediation of total petroleum hydrocarbons (TPH) in crude oil contaminated soil based on laboratory and field observations. *International Biodeterioration & Biodegradation,* 90**,** 115-122.

SUN, X., ZAHIR, Z., LYNCH, K. H. & DENNIS, J. J. 2011. An antirepressor, SrpR, is involved in transcriptional regulation of the SrpABC solvent tolerance efflux pump of Pseudomonas putida S12. *Journal of bacteriology,* 193**,** 2717-2725.

SUNAGAWA, S., MENDE, D. R., ZELLER, G., IZQUIERDO-CARRASCO, F., BERGER, S. A., KULTIMA, J. R., COELHO, L. P., ARUMUGAM, M., TAP, J. & NIELSEN, H. B. 2013. Metagenomic species profiling using universal phylogenetic marker genes. *Nature methods,* 10**,** 1196-1199.

SUTTINUN, O., LUEPROMCHAI, E. & MÜLLER, R. 2013. Cometabolism of trichloroethylene: concepts, limitations and available strategies for sustained biodegradation. *Reviews in Environmental Science and Bio/Technology,* 12**,** 99-114.

SYMONS, Z. C. & BRUCE, N. C. 2006. Bacterial pathways for degradation of nitroaromatics. *Nat. Prod. Rep.,* 23**,** 845-850.

SZALENIEC, M., HAGEL, C., MENKE, M., NOWAK, P., WITKO, M. & HEIDER, J. 2007. Kinetics and mechanism of oxygen-independent hydrocarbon hydroxylation by ethylbenzene dehydrogenase. *Biochemistry,* 46**,** 7637-7646.

SZILÁGYI, A. & ZÁVODSZKY, P. 2000. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure (London, England),* 8**,** 493-504.

TAHER, L., RINNER, O., GARG, S., SCZYRBA, A., BRUDNO, M., BATZOGLOU, S. & MORGENSTERN, B. 2003. AGenDA: homology-based gene prediction. *Bioinformatics,* 19**,** 1575-1577.

TAKAI, K., KOBAYASHI, H., NEALSON, K. H. & HORIKOSHI, K. 2003. Deferribacter desulfuricans sp. nov., a novel sulfur-, nitrate-and arsenate-reducing thermophile

isolated from a deep-sea hydrothermal vent. *International journal of systematic and evolutionary microbiology,* 53**,** 839-846.

TAKEDA, H., YAMADA, A., MIYAUCHI, K., MASAI, E. & FUKUDA, M. 2004. Characterization of transcriptional regulatory genes for biphenyl degradation in Rhodococcus sp. strain RHA1. *Journal of bacteriology,* 186**,** 2134-2146.

TAMAKI, H., WRIGHT, C. L., LI, X., LIN, Q., HWANG, C., WANG, S., THIMMAPURAM, J., KAMAGATA, Y. & LIU, W.-T. 2011. Analysis of 16S rRNA amplicon sequencing options on the Roche/454 next-generation Titanium sequencing platform. *PloS one,* 6**,** e25263.

TAMURA, K., PETERSON, D., PETERSON, N., STECHER, G., NEI, M. & KUMAR, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution,* 28**,** 2731-2739.

TANASE, A.-M., IONESCU, R., CHICIUDEAN, I., VASSU, T. & STOICA, I. 2013. Characterization of hydrocarbon< i>-</i> degrading bacterial strains isolated from oil-polluted soil. *International Biodeterioration & Biodegradation,* 84**,** 150-154.

TANG, J., WANG, M., WANG, F., SUN, Q. & ZHOU, Q. 2011. Eco-toxicity of petroleum hydrocarbon contaminated soil. *Journal of Environmental Sciences,* 23**,** 845-851.

TANG, S., ANTONOV, I. & BORODOVSKY, M. 2013. MetaGeneTack: ab initio detection of frameshifts in metagenomic sequences. *Bioinformatics,* 29**,** 114-116.

TANNER, M. A., EVERETT, C. L., COLEMAN, W. J., YANG, M. M. & YOUVAN, D. C. 2000. Complex microbial communities inhabiting sulfide-rich black mud from marine coastal environments. *Biotechnology et alia,* 8**,** 1-16.

TAYLOR, A. E., ARP, D. J., BOTTOMLEY, P. J. & SEMPRINI, L. 2010. Extending the alkene substrate range of vinyl chloride utilizing Nocardioides sp. strain JS614 with ethene oxide. *Applied microbiology and biotechnology,* 87**,** 2293-2302.

TAYLOR, A. E., VAJRALA, N., GIGUERE, A. T., GITELMAN, A. I., ARP, D. J., MYROLD, D. D., SAYAVEDRA-SOTO, L. & BOTTOMLEY, P. J. 2013. Use of aliphatic n-alkynes to discriminate soil nitrification activities of ammonia-oxidizing thaumarchaea and bacteria. *Applied and environmental microbiology,* 79**,** 6544-6551.

TEELING, H., WALDMANN, J., LOMBARDOT, T., BAUER, M. & GLÖCKNER, F. O. 2004. TETRA: a web-service and a stand-alone program for the analysis and comparison of tetranucleotide usage patterns in DNA sequences. *BMC bioinformatics,* 5**,** 163.

TEHEI, M., FRANZETTI, B., MADERN, D., GINZBURG, M., GINZBURG, B. Z., GIUDICI-ORTICONI, M.-T., BRUSCHI, M. & ZACCAI, G. 2003. Adaptation to extreme environments: macromolecular dynamics in bacteria compared in vivo by neutron scattering. *EMBO reports,* 5**,** 66-70.

TEMKAR, A. 2011. *Designing a Winogradsky Column to Simulate, Enrich and Predict Microcosm Biofilm Patterns: Building a Tool to Grow Microbes As They Are*, GRIN Verlag.

TENBRINK, F., SCHINK, B. & KRONECK, P. M. H. 2011. Exploring the active site of the tungsten, iron-sulfur enzyme acetylene hydratase. *Journal of bacteriology,* 193**,** 1229-1236.

TERAMAE, K. H., DE SOUSA, M. L., DE ANGELIS, D. D. F. & BIDOIA, E. D. 2012. Ecotoxicological Tests in Winogradsky Columns Contaminated by Used Lubricating Oil and Biodiesel. *Environmental Management and Sustainable Development,* 1**,** p139-150.

THAUER, R. K. 2011. Anaerobic oxidation of methane with sulfate: on the reversibility of the reactions that are catalyzed by enzymes also involved in methanogenesis from CO 2. *Current Opinion in Microbiology,* 14**,** 292-299.

THI NHI CONG, L. 2008. *Degradation of branched chain aliphatic and aromatic petroleum hydrocarbons by microorganisms.* Universitätsbibliothek.

THOMAS, T., GILBERT, J. & MEYER, F. 2012. Metagenomics-a guide from sampling to data analysis. *Microb Inform Exp,* 2.

THRONE-HOLST, M., WENTZEL, A., ELLINGSEN, T. E., KOTLAR, H.-K. & ZOTCHEV, S. B. 2007. Identification of novel genes involved in long-chain n-alkane degradation by Acinetobacter sp. strain DSM 17874. *Applied and environmental microbiology,* 73**,** 3327-3332.

TIAN, G., YIN, X., LUO, H., XU, X., BOLUND, L. & ZHANG, X. 2010. Sequencing bias: comparison of different protocols of microRNA library construction. *BMC biotechnology,* 10**,** 64.

TIERNEY, M. & YOUNG, L. 2010. Anaerobic degradation of aromatic hydrocarbons. *Handbook of Hydrocarbon and Lipid Microbiology.* Springer.

TIMP, W., MIRSAIDOV, U. M., WANG, D., COMER, J., AKSIMENTIEV, A. & TIMP, G. 2010. Nanopore sequencing: electrical measurements of the code of life. *Nanotechnology, IEEE Transactions on,* 9**,** 281-294.

TINBERG, C. E. & LIPPARD, S. J. 2011. Dioxygen activation in soluble methane monooxygenase. *Accounts of chemical research,* 44**,** 280-288.

TINBERG, C. E., SONG, W. J., IZZO, V. & LIPPARD, S. J. 2011. Multiple roles of component proteins in bacterial multicomponent monooxygenases: phenol hydroxylase and toluene/o-xylene monooxygenase from Pseudomonas sp. OX1. *Biochemistry,* 50**,** 1788-1798.

TODA, H., IMAE, R., KOMIO, T. & ITOH, N. 2012. Expression and characterization of styrene monooxygenases of Rhodococcus sp. ST-5 and ST-10 for synthesizing enantiopure (S)-epoxides. *Applied microbiology and biotechnology,* 96**,** 407-418.

TORRES PAZMINO, D., WINKLER, M., GLIEDER, A. & FRAAIJE, M. 2010. Monooxygenases as biocatalysts: Classification, mechanistic aspects and biotechnological applications. *Journal of biotechnology,* 146**,** 9-24.

TRESEDER, K. K., BALSER, T. C., BRADFORD, M. A., BRODIE, E. L., DUBINSKY, E. A., EVINER, V. T., HOFMOCKEL, K. S., LENNON, J. T., LEVINE, U. Y. & MACGREGOR, B. J. 2012. Integrating microbial ecology into ecosystem models: challenges and priorities. *Biogeochemistry,* 109**,** 7-18.

TRINDADE-SILVA, A. E., RUA, C. P., ANDRADE, B. G., VICENTE, A. C. P., SILVA, G. G., BERLINCK, R. G. & THOMPSON, F. L. 2013. Polyketide synthase gene diversity within the microbiome of the sponge Arenosclera brasiliensis, endemic to the Southern Atlantic Ocean. *Applied and environmental microbiology,* 79**,** 1598-1605.

TRINGE, S. G., VON MERING, C., KOBAYASHI, A., SALAMOV, A. A., CHEN, K., CHANG, H. W., PODAR, M., SHORT, J. M., MATHUR, E. J. & DETTER, J. C. 2005. Comparative metagenomics of microbial communities. *Science,* 308**,** 554-557.

TRIVEDI, S., RAO, S. R. & GEHLOT, H. S. 2005. Nucleic acid stability in thermophilic prokaryotes: a review. *J Cell Mol Biol,* 4**,** 61-69.

TURNBAUGH, P. J., HAMADY, M., YATSUNENKO, T., CANTAREL, B. L., DUNCAN, A., LEY, R. E., SOGIN, M. L., JONES, W. J., ROE, B. A. & AFFOURTIT, J. P. 2009. A core gut microbiome in obese and lean twins. *Nature,* 457**,** 480-484.

TYSON, G. W., CHAPMAN, J., HUGENHOLTZ, P., ALLEN, E. E., RAM, R. J., RICHARDSON, P. M., SOLOVYEV, V. V., RUBIN, E. M., ROKHSAR, D. S. & BANFIELD, J. F. 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature,* 428**,** 37-43.

UENO, A., ITO, Y., YUMOTO, I. & OKUYAMA, H. 2007. Isolation and characterization of bacteria from soil contaminated with diesel oil and the possible use of these in

autochthonous bioaugmentation. *World Journal of Microbiology and Biotechnology,* 23**,** 1739-1745.

UHLIK, O., STREJCEK, M., HROUDOVA, M., DEMNEROVA, K. & MACEK, T. 2013. Identification and Characterization of Bacteria with Bioremediation Potential: from Cultivation to Metagenomics. *CHEMICKE LISTY,* 107**,** 614-622.

URIBE, A., ALAM, M., JOHANSSON, O., MIDTVEDT, T. & THEODORSSON, E. 1994. Microflora modulates endocrine cells in the gastrointestinal mucosa of the rat. *Gastroenterology,* 107**,** 1259-1269.

URLACHER, V. B. & EIBEN, S. 2006. Cytochrome P450 monooxygenases: perspectives for synthetic application. *Trends in biotechnology,* 24**,** 324-330.

URLACHER, V. B. & GIRHARD, M. 2012. Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. *Trends in biotechnology,* 30**,** 26-36.

VAN BEILEN, J. B. & FUNHOFF, E. G. 2007. Alkane hydroxylases involved in microbial alkane degradation. *Applied Microbiology and Biotechnology,* 74**,** 13-21.

VAN BEILEN, J. B., FUNHOFF, E. G., VAN LOON, A., JUST, A., KAYSSER, L., BOUZA, M., HOLTACKERS, R., RÖTHLISBERGER, M., LI, Z. & WITHOLT, B. 2006. Cytochrome P450 alkane hydroxylases of the CYP153 family are common in alkane-degrading eubacteria lacking integral membrane alkane hydroxylases. *Applied and environmental microbiology,* 72**,** 59-65.

VAN BEILEN, J. B., LI, Z., DUETZ, W. A., SMITS, T. H. M. & WITHOLT, B. 2003. Diversity of Alkane Hydroxylase Systems in the Environment. *Oil & Gas Science and Technology,* 58**,** 427-440.

VAN BEILEN, J. B., PANKE, S., LUCCHINI, S., FRANCHINI, A. G., RÖTHLISBERGER, M. & WITHOLT, B. 2001. Analysis of Pseudomonas putida alkane-degradation gene clusters and flanking insertion sequences: evolution and regulation of the alk genes. *Microbiology,* 147**,** 1621-1630.

VAN BEILEN, J. B., WUBBOLTS, M. G. & WITHOLT, B. 1994. Genetics of alkane oxidation byPseudomonas oleovorans. *Biodegradation,* 5**,** 161-174.

VAN DER MEER, J. 2008. *A genomic view on the evolution of catabolic pathways and bacterial adaptation to xenobiotic compounds*, Norfolk, UK: Caister Academic Press.

VAN HYLCKAMA VLIEG, J. E., LEEMHUIS, H., SPELBERG, J. H. L. & JANSSEN, D. B. 2000. Characterization of the gene cluster involved in isoprene metabolism in Rhodococcus sp. strain AD45. *Journal of bacteriology,* 182**,** 1956-1963.

VARIN, T., LOVEJOY, C., JUNGBLUT, A. D., VINCENT, W. F. & CORBEIL, J. 2012. Metagenomic analysis of stress genes in microbial mat communities from Antarctica and the High Arctic. *Applied and environmental microbiology,* 78**,** 549-559.

VERFÜRTH, K., PIERIK, A. J., LEUTWEIN, C., ZORN, S. & HEIDER, J. 2004. Substrate specificities and electron paramagnetic resonance properties of benzylsuccinate synthases in anaerobic toluene and m-xylene metabolism. *Archives of microbiology,* 181**,** 155-162.

VERMA, V., RAJU, S. C., KAPLEY, A., KALIA, V. C., KANADE, G. S., DAGINAWALA, H. F. & PUROHIT, H. J. 2011. Degradative potential of Stenotrophomonas strain HPC383 having genes homologous to dmp operon. *Bioresource technology,* 102**,** 3227-3233.

VERNER-JEFFREYS, D. W., WELCH, T. J., SCHWARZ, T., POND, M. J., WOODWARD, M. J., HAIG, S. J., RIMMER, G. S., ROBERTS, E., MORRISON, V. & BAKER-AUSTIN, C. 2009. High prevalence of multidrug-tolerant bacteria and associated antimicrobial resistance genes isolated from ornamental fish and their carriage water. *PLoS One,* 4**,** e8388.

VIEITES, J. M., GUAZZARONI, M. E., BELOQUI, A., GOLYSHIN, P. N. & FERRER, M. 2009. Metagenomics approaches in systems microbiology. *FEMS microbiology reviews,* 33**,** 236-255.

VINAS, M., SABATÉ, J., ESPUNY, M. J. & SOLANAS, A. M. 2005. Bacterial community dynamics and polycyclic aromatic hydrocarbon degradation during bioremediation of heavily creosote-contaminated soil. *Applied and Environmental Microbiology,* 71**,** 7008-7018.

VON BERGEN, M., JEHMLICH, N., TAUBERT, M., VOGT, C., BASTIDA, F., HERBST, F.-A., SCHMIDT, F., RICHNOW, H.-H. & SEIFERT, J. 2013. Insights from quantitative metaproteomics and protein-stable isotope probing into microbial ecology. *The ISME journal*.

VON NETZER, F., PILLONI, G., KLEINDIENST, S., KRÜGER, M., KNITTEL, K., GRÜNDGER, F. & LUEDERS, T. 2013. Enhanced gene detection assays for fumarate-adding enzymes allow uncovering of anaerobic hydrocarbon degraders in terrestrial and marine systems. *Applied and environmental microbiology,* 79**,** 543-552.

WACHÉ, Y. 2013. Production of Dicarboxylic Acids and Flagrances by Yarrowia lipolytica. *Yarrowia lipolytica.* Springer.

WALKER, C., DE LA TORRE, J., KLOTZ, M., URAKAWA, H., PINEL, N., ARP, D., BROCHIER-ARMANET, C., CHAIN, P., CHAN, P. & GOLLABGIR, A. 2010. Nitrosopumilus maritimus genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proceedings of the National Academy of Sciences,* 107**,** 8818-8823.

WANG, H., LI, X., MA, Y. & SONG, J. 2014a. Characterization and high-level expression of a metagenome-derived alkaline pectate lyase in recombinant< i> Escherichia coli</i>. *Process Biochemistry,* 49**,** 69-76.

WANG, J., XU, Y., YANG, Z., HUANG, R., CHEN, J., WANG, R. & LIN, Y. 2013a. Toxicity of Carbon Nanotubes. *Current drug metabolism,* 14**,** 891-899.

WANG, L.-Y., GAO, C.-X., MBADINGA, S. M., ZHOU, L., LIU, J.-F., GU, J.-D. & MU, B.-Z. 2011. Characterization of an alkane-degrading methanogenic enrichment culture from production water of an oil reservoir after 274 days of incubation. *International Biodeterioration & Biodegradation,* 65**,** 444-450.

WANG, L.-Y., KE, W.-J., SUN, X.-B., LIU, J.-F., GU, J.-D. & MU, B.-Z. 2014b. Comparison of bacterial community in aqueous and oil phases of water-flooded petroleum reservoirs using pyrosequencing and clone library approaches. *Applied microbiology and biotechnology,* 98**,** 4209-4221.

WANG, Q., QIAN, C., ZHANG, X.-Z., LIU, N., YAN, X. & ZHOU, Z. 2012a. Characterization of a novel thermostable β-glucosidase from a metagenomic library of termite gut. *Enzyme and microbial technology*.

WANG, S., CUI, Z., ZHENG, L., GAO, W. & LI, Q. 2012b. Characterization of PAHs-degrading bacterium Porphyrobacter sp. D22F and its ecological niche in oil-degrading consortium D22-1. *Chin J Appl Environ Biol,* 18**,** 122-127.

WANG, W., MA, T., LIAN, K., ZHANG, Y., TIAN, H., JI, K. & LI, G. 2013b. Genetic Analysis of Benzothiophene Biodesulfurization Pathway of Gordonia terrae Strain C-6. *PloS one,* 8**,** e84386.

WANG, Z., FINGAS, M. & SERGY, G. 1994. Study of 22-year-old Arrow oil samples using biomarker compounds by GC/MS. *Environmental science & technology,* 28**,** 1733-1746.

WANG, Z. & FINGAS, M. F. 2003. Development of oil hydrocarbon fingerprinting and identification techniques. *Marine Pollution Bulletin,* 47**,** 423-452.

WANG, Z. & STOUT, S. 2010. *Oil spill environmental forensics: fingerprinting and source identification*, Academic Press.

WARDLAW, G. D., AREY, J. S., REDDY, C. M., NELSON, R. K., VENTURA, G. T. & VALENTINE, D. L. 2008. Disentangling oil weathering at a marine seep using GC x GC: broad metabolic specificity accompanies subsurface petroleum biodegradation. *Environmental science & technology,* 42**,** 7166.

WARNECKE, F., LUGINBÜHL, P., IVANOVA, N., GHASSEMIAN, M., RICHARDSON, T. H., STEGE, J. T., CAYOUETTE, M., MCHARDY, A. C., DJORDJEVIC, G. & ABOUSHADI, N. 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature,* 450**,** 560-565.

WARREN, R. L., SUTTON, G. G., JONES, S. J. & HOLT, R. A. 2007. Assembling millions of short DNA sequences using SSAKE. *Bioinformatics,* 23**,** 500-501.

WEELINK, S. A. B., SMIDT, H., TALARICO SAIA, F., RIJPSTRA, I., RÖLING, W., STAMS, A. J. M. & DOESBURG, V., W.C.J 2009. A strictly anaerobic betaproteobacterium Georgfuchsia toluolica gen. nov., sp. nov. degrades aromatic compounds with Fe(III), Mn(IV) or nitrate as an electron acceptor. *FEMS Microbiology Ecology,* 70**,** 575.

WEIJERS, C., JONGEJAN, H., FRANSSEN, M., DE GROOT, A. & DE BONT, J. 1995. Dithiol-and NAD-dependent degradation of epoxyalkanes by Xanthobacter Py2. *Applied microbiology and biotechnology,* 42**,** 775-781.

WENTZEL, A., ELLINGSEN, T. E., KOTLAR, H.-K., ZOTCHEV, S. B. & THRONE-HOLST, M. 2007. Bacterial metabolism of long-chain n-alkanes. *Applied microbiology and biotechnology,* 76**,** 1209-1221.

WÉRY, N., MONTEIL, C., POURCHER, A.-M. & GODON, J.-J. 2010. Human-specific fecal bacteria in wastewater treatment plant effluents. *Water research,* 44**,** 1873-1883.

WHITMAN, W. B. & PARTE, A. C. 2009. *Systematic bacteriology*, Springer.

WHYTE, L., SMITS, T., LABBE, D., WITHOLT, B., GREER, C. & VAN BEILEN, J. 2002a. Gene cloning and characterization of multiple alkane hydroxylase systems in Rhodococcus strains Q15 and NRRL B-16531. *Applied and environmental microbiology,* 68**,** 5933-5942.

WHYTE, L. G., SMITS, T. H. M., LABBE, D., WITHOLT, B., GREER, C. W. & VAN BEILEN, J. B. 2002b. Gene Cloning and Characterization of Multiple Alkane Hydroxylase Systems in Rhodococcus Strains Q15 and NRRL B-16531. *Applied and Environmental Microbiology,* 68**,** 5933-5942.

WIDDEL, F., KNITTEL, K. & GALUSHKO, A. 2010a. Anaerobic hydrocarbon-degrading microorganisms: an overview. *Handbook of hydrocarbon and lipid microbiology.* Springer.

WIDDEL, F., KNITTEL, K. & GALUSHKO, A. 2010b. *Anaerobic Hydrocarbon-Degrading Microorganisms: An Overview*, Berlin, Heidelberg, Springer Berlin Heidelberg.

WIDDEL, F. & RABUS, R. 2001. Anaerobic biodegradation of saturated and aromatic hydrocarbons. *Current opinion in biotechnology,* 12**,** 259–276.

WIDERSTEN, M., GURELL, A. & LINDBERG, D. 2010. Structure–function relationships of epoxide hydrolases and their potential use in biocatalysis. *Biochimica et Biophysica Acta (BBA)-General Subjects,* 1800**,** 316-326.

WIEGEL, J. 2006. The Genus Xanthobacter. *The Prokaryotes.* Springer.

WIEHE, T., GEBAUER-JUNG, S., MITCHELL-OLDS, T. & GUIGÓ, R. 2001. SGP-1: prediction and validation of homologous genes based on sequence alignments. *Genome Research,* 11**,** 1574-1583.

WIKOFF, W. R., ANFORA, A. T., LIU, J., SCHULTZ, P. G., LESLEY, S. A., PETERS, E. C. & SIUZDAK, G. 2009. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proceedings of the National Academy of Sciences,* 106**,** 3698-3703.

WILBON, P. A., CHU, F. & TANG, C. 2013. Progress in renewable polymers from natural terpenes, terpenoids, and rosin. *Macromolecular rapid communications,* 34**,** 8-37.

WILKE, A., GLASS, E. M., BISCHOF, J., BRAITHWAITE, D., D'SOUZA, M., GERLACH, W., HARRISON, T., KEEGAN, K., MATTHEWS, H. & PACZIAN, T. 2013. MG-RAST Technical Report and Manual for version 3.3. 6, revision 2.

WILKES, H., RABUS, R., FISCHER, T., ARMSTROFF, A., BEHRENDS, A. & WIDDEL, F. 2002. Anaerobic degradation of n-hexane in a denitrifying bacterium: further degradation of the initial intermediate (1-methylpentyl) succinate via C-skeleton rearrangement. *Archives of microbiology,* 177**,** 235-243.

WILLIAMS, G. J. 2013. Engineering polyketide synthases and nonribosomal peptide synthetases. *Current opinion in structural biology,* 23**,** 603-612.

WILMS, R., SASS, H., KÖPKE, B., CYPIONKA, H. & ENGELEN, B. 2007. Methane and sulfate profiles within the subsurface of a tidal flat are reflected by the distribution of sulfate-reducing bacteria and methanogenic archaea. *FEMS microbiology ecology,* 59**,** 611-621.

WILMS, R., SASS, H., KÖPKE, B., KÖSTER, J., CYPIONKA, H. & ENGELEN, B. 2006. Specific bacterial, archaeal, and eukaryotic communities in tidal-flat sediments along a vertical profile of several meters. *Applied and Environmental Microbiology,* 72**,** 2756-2764.

WILSON, M. C. & PIEL, J. 2013. Metagenomic Approaches for Exploiting Uncultivated Bacteria as a Resource for Novel Biosynthetic Enzymology. *Chemistry & biology,* 20**,** 636-647.

WITTCOFF, H. A., REUBEN, B. G. & PLOTKIN, J. S. 2012. *Industrial organic chemicals,* John Wiley & Sons.

WOOD, T. K. 2008. Molecular approaches in bioremediation. *Current Opinion in Biotechnology,* 19**,** 572-578.

WOODHOUSE, J. N., FAN, L., BROWN, M. V., THOMAS, T. & NEILAN, B. A. 2013. Deep sequencing of non-ribosomal peptide synthetases and polyketide synthases from the microbiomes of Australian marine sponges. *The ISME journal*.

WOOLEY, J. C. & YE, Y. 2010. Metagenomics: facts and artifacts, and computational challenges. *Journal of computer science and technology,* 25**,** 71-81.

WU, D., DOROUD, L. & EISEN, J. A. 2013. TreeOTU: Operational Taxonomic Unit Classification Based on Phylogenetic Trees. *arXiv preprint arXiv:1308.6333*.

WU, M., ETTWIG, K., JETTEN, M., STROUS, M., KELTJENS, J. & VAN NIFTRIK, L. 2011. A new intra-aerobic metabolism in the nitrite-dependent anaerobic methane-oxidizing bacterium Candidatus' Methylomirabilis oxyfera'. *Biochemical Society Transactions,* 39**,** 243.

XIA, X. 2013. DAMBE5: A comprehensive software package for data analysis in molecular biology and evolution. *Molecular biology and evolution,* 30**,** 1720-1728.

XIANQING, L., DUJIE, H., YOUJUN, T., GUOYI, H. & BO, X. 2003. Molecular geochemical evidence for the origin of natural gas from dissolved hydrocarbon in Ordovician formation waters in central Ordos Basin. *Chinese Journal of Geochemistry,* 22**,** 193-202.

XIE, M., ALONSO, H. & ROUJEINIKOVA, A. 2011. An improved procedure for the purification of catalytically active alkane hydroxylase from Pseudomonas putida GPo1. *Applied biochemistry and biotechnology,* 165**,** 823-831.

XIE, W., WANG, F., GUO, L., CHEN, Z., SIEVERT, S. M., MENG, J., HUANG, G., LI, Y., YAN, Q. & WU, S. 2010. Comparative metagenomics of microbial communities

inhabiting deep-sea hydrothermal vent chimneys with contrasting chemistries. *The ISME journal,* 5**,** 414-426.

XIONG, Z.-Q., WANG, J.-F., HAO, Y.-Y. & WANG, Y. 2013. Recent Advances in the Discovery and Development of Marine Microbial Natural Products. *Marine drugs,* 11**,** 700-717.

XIONG, Z. 2013. Metagenomic-Guided Antibiotics Discovery. *Clin Microbial,* 2**,** 101.

YAKIMOV, M. M., TIMMIS, K. N. & GOLYSHIN, P. N. 2007. Obligate oil-degrading marine bacteria. *Current opinion in biotechnology,* 18**,** 257-266.

YAMADA, M., NAKASONE, K., TAMEGAI, H., KATO, C., USAMI, R. & HORIKOSHI, K. 2000. Pressure Regulation of Soluble Cytochromesc in a Deep-Sea Piezophilic Bacterium, Shewanella violacea. *Journal of bacteriology,* 182**,** 2945-2952.

YAN, J., WANG, L., FU, P. P. & YU, H. 2004. Photomutagenicity of 16 polycyclic aromatic hydrocarbons from the US EPA priority pollutant list. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis,* 557**,** 99-108.

YANG, B., DAI, Z., DING, S.-Y. & WYMAN, C. E. 2011. Enzymatic hydrolysis of cellulosic biomass. *Biofuels,* 2**,** 421-450.

YANG, C., WANG, Z., HOLLEBONE, B., BROWN, C., LANDRIAULT, M., FIELDHOUSE, B. & YANG, Z. 2012. Application of Light Petroleum Biomarkers for Forensic Characterization and Source Identification of Spilled Light Refined Oils. *Environmental Forensics,* 13**,** 298-311.

YANG, Y. & YOOSEPH, S. 2013. SPA: a short peptide assembler for metagenomic data. *Nucleic acids research,* 41**,** e91-e91.

YE, L. & ZHANG, T. 2013. Bacterial communities in different sections of a municipal wastewater treatment plant revealed by 16S rDNA 454 pyrosequencing. *Applied Microbiology and Biotechnology***,** 1-10.

YE, Y., WEI, B., WEN, L. & RAYNER, S. 2013. BlastGraph: a comparative genomics tool based on BLAST and graph algorithms. *Bioinformatics,* 29**,** 3222-3224.

YEN, K.-M., SERDAR, C. M. & GUNSALUS, I. C. 1988. Genetics of naphthalene catabolism in pseudomonads. *Critical reviews in microbiology,* 15**,** 247-268.

YERGEAU, E., SANSCHAGRIN, S., BEAUMIER, D. & GREER, C. W. 2012a. Metagenomic analysis of the bioremediation of diesel-contaminated Canadian high Arctic soils. *PloS one,* 7**,** e30058.

YERGEAU, E., SANSCHAGRIN, S., BEAUMIER, D. & GREER, C. W. 2012b. Metagenomic analysis of the bioremediation of diesel-contaminated Canadian high arctic soils. *PloS one,* 7**,** 1-10.

YOK, N. G. & ROSEN, G. L. 2011. Combining gene prediction methods to improve metagenomic gene annotation. *BMC bioinformatics,* 12**,** 20.

YOON, S. 2015. COMPOSITION, METHOD, AND APPARATUS FOR CRUDE OIL REMEDIATION. US Patent 20,150,014,221.

ZAKARIA, M. P., TAKADA, H., TSUTSUMI, S., OHNO, K., YAMADA, J., KOUNO, E. & KUMATA, H. 2002. Distribution of polycyclic aromatic hydrocarbons (PAHs) in rivers and estuaries in Malaysia: a widespread input of petrogenic PAHs. *Environmental science & technology,* 36**,** 1907-1918.

ZAKRZEWSKI, M., BEKEL, T., ANDER, C., PÜHLER, A., RUPP, O., STOYE, J., SCHLÜTER, A. & GOESMANN, A. 2012. MetaSAMS–A novel software platform for taxonomic classification, functional annotation and comparative analysis of metagenome datasets. *Journal of biotechnology*.

ZARDINI, H. Z., DAVARPANAH, M., SHANBEDI, M., AMIRI, A., MAGHREBI, M. & EBRAHIMI, L. 2013. Microbial toxicity of ethanolamines—Multiwalled carbon nanotubes. *Journal of Biomedical Materials Research Part A.*

ZEINALI, M., VOSSOUGHI, M. & ARDESTANI, S. K. 2008. Naphthalene metabolism in Nocardia otitidiscaviarum strain TSH1, a moderately thermophilic microorganism. *Chemosphere,* 72**,** 905-909.

ZEMAN, N. R. 2012. Thermally enhanced bioremediation of LNAPL. *Dynamic Biochemistry, Process Biotechnology and Molecular Biology, at Global Science Books* 48-56.

ZENGLER, K., HEIDER, J., ROSSELLÓ-MORA, R. & WIDDEL, F. 1999. Phototrophic utilization of toluene under anoxic conditions by a new strain of Blastochloris sulfoviridis. *Archives of Microbiology,* 172**,** 204-212.

ZERBINO, D. R. & BIRNEY, E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome research,* 18**,** 821-829.

ZEYER, J., KOCHER, H. & TIMMIS, K. 1986. Influence of para-substituents on the oxidative metabolism of o-nitrophenols by Pseudomonas putida B2. *Applied and environmental microbiology,* 52**,** 334-339.

ZHANG, A., ZHAO, R., JIN, P., MA, L., XIONG, X., XIE, T., PEI, X., YU, L., YIN, X. & WANG, Q. 2013a. Discovery of a novel esterase subfamily sharing an identified arm

sequence (ArmEst) by gene-specific metagenomic PCR. *Biotechnology letters,* 35**,** 1937-1944.

ZHANG, F., SHE, Y.-H., CHAI, L.-J., BANAT, I. M., ZHANG, X.-T., SHU, F.-C., WANG, Z.-L., YU, L.-J. & HOU, D.-J. 2012. Microbial diversity in long-term water-flooded oil reservoirs with different in situ temperatures in China. *Scientific reports,* 2.

ZHANG, X., QU, Y., MA, Q., ZHOU, H., LI, X., KONG, C. & ZHOU, J. 2013b. Cloning and expression of naphthalene dioxygenase genes from Comamonas sp. MQ for indigoids production. *Process Biochemistry,* 48**,** 581-587.

ZHANG, Y., TANG, X.-J., SHEN, B., YU, X.-J., WANG, E.-T. & YUAN, H.-L. 2013c. Identification and Characterization of the Butane-Utilizing Bacterium, Arthrobacter sp. PG-3-2, Harboring a Novel bmoX Gene. *Geomicrobiology Journal,* 30**,** 85-92.

ZHAO, B., WANG, H., LI, R. & MAO, X. 2010. Thalassospira xianhensis sp. nov., a polycyclic aromatic hydrocarbon-degrading marine bacterium. *International journal of systematic and evolutionary microbiology,* 60**,** 1125-1129.

ZHAO, D., LIU, C., LIU, L., ZHANG, Y., LIU, Q. & WU, W.-M. 2011. Selection of functional consortium for crude oil-contaminated soil remediation. *International Biodeterioration & Biodegradation,* 65**,** 1244-1248.

ZHONG, W., GU, T., WANG, W., ZHANG, B., LIN, X., HUANG, Q. & SHEN, W. 2010. The effects of mineral fertilizer and organic manure on soil microbial community and diversity. *Plant and soil,* 326**,** 511-522.

ZHU, L., WU, Q., DAI, J., ZHANG, S. & WEI, F. 2011. Evidence of cellulose metabolism by the giant panda gut microbiome. *Proceedings of the National Academy of Sciences,* 108**,** 17714-17719.

ZIMIN, A. V., MARÇAIS, G., PUIU, D., ROBERTS, M., SALZBERG, S. L. & YORKE, J. A. 2013. The MaSuRCA genome assembler. *Bioinformatics,* 29**,** 2669-2677.

ZYLSTRA, G. J. & GIBSON, D. T. 1991. Aromatic hydrocarbon degradation: a molecular approach. *Genetic engineering.* Springer.