Microbial Adaptation for the Degradation of Metaldehyde

Víctor Manuel Castro Gutiérrez

Doctor of Philosophy

University of York

Biology

October 2020

Abstract

Metaldehyde is used as molluscicide of choice in agriculture and horticulture, but its detection in drinking water has become a major cause of concern. Eight new metaldehyde-degrading bacterial strains were isolated from soil and comparative genomics were used to identify a highly-conserved gene cluster responsible for metaldehyde degradation in most of the isolates. Accelerated biodegradation of metaldehyde was demonstrated in multiple soils. It was confirmed that the metaldehyde-degrading population proliferates in response to the pesticide. Third generation sequencing was used to obtain a reference genome for the model degrader *Acinetobacter calcoaceticus* E1. IS91 and IS6-family insertion sequences were found surrounding the degrading gene cluster. They were hypothesised to be responsible for degrading gene dissemination to different taxa. A plasmid denominated pAME76 was found to harbour the degrading operon and bioinformatic analyses suggested that the genes are also located in plasmids in most of the other isolates.

A sensitive probe-based qPCR assay for degrading gene cluster detection was developed and the technique was successfully used in different soils. Higher number of gene copies correlated with periods of metaldehyde removal for half of the soils. The gene cluster was not detected in the other half of the soils indicating that other biological degrading pathways are also important.

The metaldehyde-degrading bacterial library was screened using batch assays to identify the best strains for removal of the compound at micropollutant concentrations from water. *A. calcoaceticus* E1 and *Sphingobium* CMET-H were used for bioaugmentation of pilot-scale slow sand filters. The former was not able to achieve metaldehyde-compliant water even though a remarkably high population was present in the sand. In contrast, the latter could swiftly bring metaldehyde concentrations to compliant levels that persisted for 14 d until reactor shutdown, even though its population in the top layer of the sand filters was relatively low.

Table of Contents

List of Figures

List of Tables

Acknowledgements

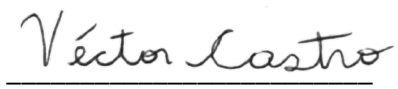
First, I would like to thank my Dad, who, through his example, taught me to always do my best, no matter the obstacles. To my wife Diana for sharing these years of my life, for making me enjoy the highs and supporting me greatly during the lows. To Mum, who taught me that what matters is the inner value of people. Adri, Pablo, Javier, and Caro for all their help from the start of this process up until the very end, I always felt your support. Eu, for making me enjoy the trips back to Costa Rica, you were the connection to my homeland. Leti who believed and supported us all along.

I am profoundly grateful to Professor James Moir, who guided me and supported me throughout the PhD. I could not have had a better supervisor. Special thanks to Dr. Thorunn Helgason who helped me every time I needed advice. I am also very grateful to Dr. Francis Hassard, who was vital in expanding the magnitude and impact of the project and making me feel more than welcome in Cranfield. Thanks to my TAP panel, Dr. Ville Friman and Dr. Daniel Jeffares, for all their helpful and knowledgeable suggestions.

I would especially like to thank Dr. Carlos Rodríguez, I could not have made it this far without him. I would also like to thank my friends at CICA and the University of Costa Rica for their support. Finally, my lab mates at L1, especially Ed, Claire, Liam and Emily who made me enjoy this stage of my life in York.

Author's Declaration

I declare that the work presented in this thesis is my own original work, unless explicitly stated otherwise. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.



Víctor Manuel Castro Gutiérrez

October 2020

Chapter 1 Section 1.4 "Detection and Quantification of Pesticide-degrading Genes in the Environment" has been approved for publication as part of the review chapter:
Fuller, E., Castro-Gutiérrez, V., Cambronero-Heinrichs, J.C., Rodriguez-Rodríguez, C.E. 2020. Using Molecular Methods to Identify and Monitor Xenobiotic-degrading Genes for Bioremediation. In: Bidoia, E., Montagnolli, R. Biodegradation, Pollutants and Bioremediation Principles. Taylor and Francis Group LLC.

A modified version of **Chapter 2** has been published as a peer-reviewed journal paper:
Castro-Gutiérrez, V., Fuller, E., Thomas, J.C., Sinclair, C.J., Johnson, S., Helgason, T., Moir, J.W.B., 2020. Genomic basis for pesticide degradation revealed by selection, isolation and characterisation of a library of metaldehyde-degrading strains from soil. Soil Biology and Biochemistry 142. doi:10.1016/j.soilbio.2019.107702.
The published version of the manuscript has been included as Annex 1.

A manuscript based on **Chapter 3** is currently being prepared for peer-review and publication.

The work presented in **Chapter 4** was possible due to a collaboration with researchers from Cranfield University (Laura Pickering, PhD student, and Francis Hassard, Lecturer in Public Health Microbiology). A three-month placement at Cranfield University was undertaken. Cranfield University provided metaldehyde reference material, equipment, reagents and culture media for growth and inoculation of strains, equipment and reagents for flow cytometry, sand and raw water for all assays. Metaldehyde quantification by LC/MS-MS was performed at Cranfield University with equipment and technical support of their technical staff. Installation and operation of slow sand filters, and physicochemical analyses of influent and effluent water were carried out by Cranfield University's researchers. The original experimental design of this chapter was my own work and was slightly modified with input from Cranfield University's researchers for improvement and practical considerations. The chapter is being prepared for peer-review and publication. The data presented in this chapter are subject to a client confidentiality agreement. Technical performance characterisation of the slow sand filters will feature in a Cranfield University doctoral thesis authored by Laura Pickering.

CHAPTER 1: GENERAL INTRODUCTION[1]

1.1 The Environmental behaviour of pesticides

Pesticides are generally classified as herbicides, insecticides, acaricides, nematicides, fungicides or molluscicides. Even though the use of these compounds has brought numerous benefits, such as the improvement on agricultural production yields and the elimination of pathogen-associated vectors, migration of pesticides to locations other than the ones originally intended leads to economic losses generated by inefficient pest control, toxic effects and environmental pollution (Waite et al. 2002). The two fundamental processes that determine the environmental fate of a pesticide are its mobility and its persistence (Kerle et al. 2007).

1.1.1 Mobility

Pesticides undergo redistribution from the initial application point. Three main characteristics have a major influence on pesticide mobility: adsorption capacity, water solubility and vaporization potential (Kerle et al. 2007).

Adsorption refers to the attraction between a chemical compound and soil, vegetation or other surfaces, being a chemical or physical process in which the molecules of a fluid come into contact with a surface and are retained for a specific amount of time (Navarro et al. 2007). The strength of this process depends on the physical and chemical properties of both the adsorptive surface and the pesticide (Pignatello and Xing 1996; Senesi 1992). Adsorption determines to a great extent the processes of pesticide volatilization, leaching, runoff, biodegradation and soil extraction by plants, as it is fundamental in determining pesticide concentrations in aqueous and gaseous phases in soil, and the abovementioned processes occur

[1] Chapter 1 Section 1.4 "Detection and Quantification of Pesticide-degrading Genes in the Environment" has been approved for publication as part of the review chapter:
Fuller, E., Castro-Gutiérrez, V., Cambronero-Heinichs, J.C., Rodriguez-Rodríguez, C.E. 2020. Using Molecular Methods to Identify and Monitor Xenobiotic-degrading Genes for Bioremediation. In: Bidoia, E., Montagnolli, R. Biodegradation, Pollutants and Bioremediation Principles. Taylor and Francis Group LLC.

in or through these two fluid phases (Candela 2003). This way, pesticides that are moderately adsorbed to soil particles tend to remain in the zone surrounding the plant roots, where they can be available for assimilation by plants and chemical or microbial degradation; on the contrary, pesticides that are strongly adsorbed to soil particles are usually much less available for microbial degradation and plant assimilation. Those compounds that are only weakly adsorbed are more prone to move through soil along with infiltration water, in the process called leaching, or with runoff water and they eventually have the potential of contaminating groundwater or surface water sources (Kerle et al. 2007).

The distribution of organic molecules (pesticides) between organic phases and water can be quantified by different partition coefficients. The organic phase can be octanol (K_{ow}), soil (K_d) or soil normalized by organic carbon content (K_{oc}). Hydrophobic organic molecules typically exhibit decreasing polarity, water solubility and increasing size. The more hydrophobic the contaminant, the more likely it is to partition into the organic phase. The partition coefficient provides a fairly accurate understanding of the sorptive process occurring between water and the soil, more specifically, the soil organic matter (Piwoni and Keeley 1990).

The related concept of the solubility of a pesticide describes its propensity to dissolve in water and also affects its leaching potential. When water infiltrates downwards through soil, it carries soluble chemical compounds. Pesticides with a solubility above 30 mg L^{-1} have a greater capability to undergo leaching to groundwater bodies; furthermore, they have a greater potential to be carried by runoff on the surface (Geyikci 2011). Runoff and associated erosion have been regarded as the main transfer mechanism of pesticides from agricultural plots to aquatic ecosystems (Brady et al. 2006; Leu et al. 2004).

Pesticide vaporization is defined as the process through which the compound shifts from a solid or liquid state to a gaseous state (García 1997). This way it moves further away from the initial application spot and can later undergo photochemical degradation, can be deposited again in the soil through dew, rain or dust, and this way it can be distributed inside or outside of

the target area. The percentage of pesticide losses due to vaporization can go from less than 1% to more than 90% in the first 24 hours after pesticide application (García 1997).

1.1.2 Persistence

Persistence is another one of the characteristics that determines the environmental fate of pesticides, and it is defined as the tendency of a specific compound to maintain its molecular and chemical integrity, and its physical and functional characteristics in the medium through which it is transported and distributed after being released into the environment (Navarro et al. 2007). The persistence of a pesticide can be expressed in terms of half-life, which is determined by the resistance of a compound to degradation processes through time (García 1997). Theoretically, pesticide degradation follows an exponential reduction in its concentration that can be described by the following first-order equation (Beulke et al. 2000):

$$c = c_0 e^{-kt}$$

Where *c* stands for the concentration of pesticide (in mg kg^{-1} of soil) at time *t* (in days), c_0 is the initial pesticide concentration (in mg kg^{-1} of soil) and *k* is the daily degradation rate. This way, half-life can be expressed by the equation:

$$t_{1/2} = \frac{ln2}{k}$$

Final aerobic degradation products of any organic compound are water, carbon dioxide and minerals (Kerle et al. 2007). Nevertheless, it is important to point out that intermediate degradation products of metabolites derived from a pesticide have different persistence, toxic properties and mobility than the original compound, that can be either greater or smaller (Candela 2003).

A pesticide can be degraded through the action of sunlight, by chemical degradation or by biological degradation. Direct **photodegradation** occurs when a pesticide molecule receives energy directly from ultraviolet radiation coming from the sun, which excites the molecules so that these break apart or form less stable chemical bonds. Also, the term applies to an indirect process in which sunlight energy is first absorbed by other compounds and these later transmit this energy to the pesticide molecule or generate reactive chemical species different from the original compound (Navarro et al. 2007). Nonetheless, it is considered that the importance of this form of degradation is limited when it comes to pesticides.

Chemical degradation occurs through processes in which living organisms are not involved. Hydrolysis, oxidation-reduction, substitution, elimination and dehalogenation reactions, without the influence of biological activities are the ones most commonly involved in this type of degradation. Hydrolysis is one of the most important reactions for pesticide degradation that takes place in soil and water (Candela 2003; García 1997).

Biodegradation represents the main route for pesticide degradation once the compounds have been deposited in soil. It is carried out mainly by microorganisms and plants (Van Eerd et al. 2003). Due to the importance of microbial pesticide degradation in this research, a separate section will be dedicated to this topic.

1.2 Biodegradation of pesticides

Microbial biodegradation is the main mechanism that prevents accumulation of pesticides in the ecosystem (Arbeli and Fuentes 2007). Biodegradation is defined as the decomposition of a substance catalysed by enzymes *in vitro* or *in vivo*. Depending on its extent, the type of biodegradation can be classified in the following categories (Scow, 1983):

- Primary biodegradation. Any biologically induced structural transformation in the original compound that changes its molecular integrity; also called biotransformation.

- Ultimate biodegradation. The biological conversion of an organic compound to inorganic compounds and products associated with normal metabolic processes; also called mineralization.

Microbial pesticide biodegradation can occur following two routes: a) growth-associated metabolism, in which, catabolic organisms proliferate during degradation using the pesticide as a source of energy, carbon and/or nitrogen; and b) cometabolism, which is defined as the ability of a microorganism to degrade a contaminant without using it as a substrate for growth, while using a different substrate for this purpose. In this case, the contaminant is not used as a source of energy, and so, its biodegradation does not represent an advantage for the microorganism and it is a result of a fortuitous event (Nzila 2013). Therefore, during cometabolism, the degradation is in most cases mediated by non-specific reactions (Seffernick and Wackett 2001).

Bacteria and fungi represent the dominant organisms in soil regarding biomass and metabolic activity (Aldén et al. 2001); furthermore, these are considered the main groups with respect to pesticide degradation (De Schrijver and De Mot 1999). Oftentimes, fungi are in charge of biotransforming pesticides through minor structural changes in the molecule, rendering it non-toxic; and then the biotransformed pesticide is released into soil, where it is susceptible to further degradation by bacteria (Gianfreda and Rao 2004). It is important to point out that in nature, pesticide mineralization can often be carried out by the combined action of microbial communities and abiotic factors, instead of only isolated microbial species (Galiulin et al. 2001).

According to Harms et al. (2011), saprophytic fungi produce a wide array of extracellular enzymes, essential for the degradation of plant material, and which often allow the fungal degradation of organic contaminants. Because of their low substrate specificity, these enzymes allow these organisms to cometabolize structurally diverse compounds, representing different classes of contaminants. These characteristics, along with the ability of fungi to form extended

mycelial networks, allow these organisms to be efficient for different xenobiotic bioremediation purposes.

On the other hand, bacteria are more susceptible of undergoing genetic modifications than other types of microorganisms, which gives them an additional potential to increase their biodegradation capabilities (Ortiz-Hernández et al. 2011). Short generation times and high genomic plasticity allow bacteria to obtain new genes through evolution at a relatively high rate, moreover they can degrade almost any organic compound (Johnson and Spain 2003). Bacteria are also easier to grow in laboratory media than other microorganisms and are therefore easier to study.

Several mechanisms for the degradation of specific pesticides carried out by bacterial strains have been revealed and compiled in scientific literature; these include organochlorides, organophosphorus pesticides, carbamates, pyrethroids, triazines, chlorphenoxyacetic and phenylurea pesticides among others (Diez 2010; Ortiz-Hernández et al. 2011; Van Eerd et al. 2003).

An important number of pure bacterial isolates, with the ability of using specific pesticides as a sole source of carbon, nitrogen or phosphorous have been listed in several reviews (Ortiz-Hernández et al. 2011; B. Singh et al. 1999). On other occasions, mixed bacterial cultures (consortia) with pesticide degradation capacity have been obtained (De Souza et al. 1998; Mandelbaum et al. 1993; Roberts et al. 1993; Shelton and Somich 1988). Oftentimes, individual species in pure culture, isolated from these mixed communities have been unable to use the specific pesticide as a source of energy, denoting the abovementioned interspecies dependence in pesticide degradation.

1.2.1 Microbial adaptation to pesticide degradation

The abundance of microbes capable of pesticide degradation in a given setting and the efficiency of the pathway they perform in the specific environmental conditions are central to

the extent of their degradation potential; both of these factors will depend on the on the microbial population's history of exposure to the substrate (Krutz et al. 2010)

According to Arbeli and Fuentes (2007) microbial adaptation to degrade a specific chemical in soil may reflect: (1) increased degrading gene expression leading to an increase in degradation by a limited number of initial microorganisms; (2) growing numbers of degraders due to microbial growth or lateral gene transfer; (3) migration of degrading organisms from other sites and (4) the evolution of the genes and enzymes required for use of the compound as an energy source.

This is the conceptual basis of accelerated pesticide biodegradation. The term is applied to cases in which one or more previous applications of a pesticide cause a significant increase in the rate of biodegradation of such compound in a subsequent treatment due to microbial adaptation to pesticide influx. The phenomenon has been documented extensively and it may also be applied if this change is generated by a previous application of a different pesticide with a similar chemical structure (Arbeli and Fuentes 2007).

1.3 Metaldehyde

The pesticide metaldehyde is a molluscicide widely used to control snail and slug populations that can damage agricultural crops and domestic gardens (Eckert et al. 2012). It is a tetramer of acetaldehyde with the chemical formula $(CH_3CHO)_4$ and has a polar nature. Its skeletal structure is shown in Figure 1-1. Its general physical and chemical properties are shown in Table 1-1. It is moderately soluble in water and it is hydrolytically and photolytically stable (Carpenter 1989; Kegley et al. 2016).

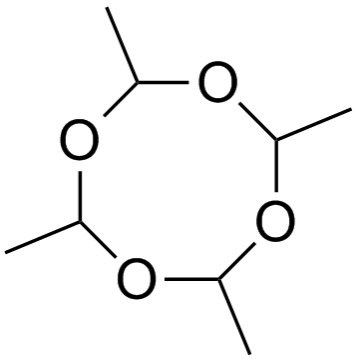


Figure 1-1. Skeletal structure of metaldehyde

Table 1-1. Chemical and physical properties of metaldehyde (PPDB 2020)

Molecular mass	176.21 g mol^{-1}
IUPAC name	2,4,6,8-tetramethyl-1,3,5,7-tetraoxocane
Boiling point	191°C (sublimes)
Water solubility at 20°C	188 mg L^{-1}
Solubility in methanol at 20°C	1730 mg L^{-1}
Vapour pressure at 20°C	6600 mPa
log organic-carbon/water partition coefficient (K_{oc})	2.38
log octanol/water partition coefficient (K_{ow}) at 20°C	0.12

Metaldehyde is a xenobiotic, i.e. not found in nature, and is synthesized from acetaldehyde via polymerization in an acidic condition at -10°C (Kekulé and Zincke 1872). The reaction results in a low yield for metaldehyde (around 8%), while paraldehyde makes up around 80% of the products (Lonza patent US3403168). Once used as a solid fuel for camping (Miller 1928), it has been used in all sectors (agriculture, horticulture, recreational land and gardens) since the 1970s as a pesticide.

Metaldehyde is applied as 1.5, 3.0 or 4.0% by weight pellets (Castle et al. 2017). In the UK, metaldehyde is used in crops such as oilseed rape, wheat, potatoes, among others (Garthwaite et al. 2018). It is normally applied as pelleted bran bait in the autumn, when the soil

is wet, molluscs prosper due to the humid weather conditions, and rain is highly likely (Green, 1996).

Approximately 135,684 kg of metaldehyde were applied in Great Britain in 2016 (Fera Science 2020). Metaldehyde (86% by area treated), ferric phosphate (13%) and methiocarb 1% were the only molluscicides used (Garthwaite et al. 2018).

Its mode of action is based on the destruction of mucocytes in molluscs, generating the production of excess mucus or slime, which ultimately leads to water loss, dehydration and death in target pests (Cragg and Vincent 1952; Triebskorn et al. 1998). Metaldehyde has a mild toxicity in humans, with ingestion of several mg kg^{-1} needed for minor effects, while severe symptoms occur at doses above 100 mg kg^{-1} and death likely at 400 mg kg^{-1} concentrations (Ellehnorn and Barceloux 1997).

Metaldehyde pellets are palatable and easily ingested by animals. After chocolate, metaldehyde is the second most common cause of poisoning in canines (Cope et al. 2006) A total of 772 cases concerning suspected metaldehyde slug bait ingestion in dogs were reported on the Veterinary Poisons Information Service database between 1985 and 2010 (Bates et al. 2012). In cats and dogs LD_{50} estimates of 207 and 500 mg kg^{-1} respectively have been reported (Gupta 2018). Birds have also been victims of metaldehyde poisoning (Reece et al. 1985).

A ban on the outdoor use of metaldehyde was planned to be introduced in Great Britain from spring 2020 (DEFRA 2018) based on an “unacceptable risk to birds and mammals”. Nonetheless, the ban was overturned in 2019 by a legal challenge from one of the producing companies (NFU 2019), metaldehyde pellets returned to the UK market and are currently awaiting a process of re-registration to determine their fate for future use.

1.3.1 Metaldehyde as a contaminant in water sources

Inadequate pesticide handling can cause the contamination of superficial water bodies or groundwater. Pesticides can enter water bodies through point or diffuse sources (Carter

2000). Point source contamination arises as a consequence of localized situations where pesticides enter a water body at a restricted number of locations; for example: commercial product spills during filling processes, leaks on spraying equipment, residual volumes on pumps and tanks, among others. On the other hand, influx through diffuse sources occurs as a consequence of the application of pesticides on the field in agricultural practice. In the case of metaldehyde, pollution of water bodies occurs primarily through this latter route (Dolan et al. 2012).

Metaldehyde pellets applied to fields are broken up by weathering and the compound is transported to water systems through pathways that include mainly surface runoff, drains and groundwater flow. Due to its low K_{oc} value (Table 1-1) it is expected that metaldehyde would be mobilized preferentially through surface runoff water in a dissolved form, instead of being transported via soil particle erosion to water systems (Asfaw et al. 2018). Due to other properties such as its moderate solubility, and low K_{ow} (Table 1-1) the compound runs off readily from fields after rainfall and can reach water bodies used for drinking water abstraction (Castle et al. 2017). In fact, it has been shown that the transfer from field to watercourse can peak at 1-4 days after rainfall (Lazartigues et al. 2013).

The European Union Drinking Water Directive (Council Directive 98/83/EC; DWD) defines the regulatory limits for individual and total pesticides in drinking water at 0.1 µg L^{-1} and 0.5 µg L^{-1} respectively. Nonetheless, it was not until 2007 that an adequate method for the determination of metaldehyde in water at very low concentrations was developed by Bristol Water and numerous breaches were detected. It was subsequently shown that the compound was present in groundwater, surface water and processed water (Stuart et al. 2011). For instance, more than 90% of the failures to comply with these limits in England in 2014 and 2015 were caused by metaldehyde (Chief Inspector of Drinking Water 2017a). Even though the levels detected do not represent an immediate threat to human health, they do present a risk to drinking water quality compliance.

Best practice methods have been promoted to prevent the migration of the compound to water courses by several groups including Metaldehyde Stewardship Group (Get Pelletwise), the Environment Agency and Natural England, among others, looking at innovative catchment and stakeholder management initiatives. The maximum recommended dose of metaldehyde in the UK is currently 210 g active substance/ha (from 1st August to 31st December) and a maximum total dose rate: 700 g active substance/ha for every calendar year; also no pellets are to be allowed to fall within a minimum of 10 metres of any field boundary or watercourse, and the product should not be applied if there is a forecast of heavy rain (Metaldehyde Stewardship Group 2020).

Moreover, farmers have been encouraged in some instances to switch to other molluscicides such as ferric phosphate (Severn Trent Water 2014); however, due to its extended use and perceived low cost, metaldehyde is still the preferred option for farmers. Moreover, slugs remain above the ground upon death with metaldehyde, whilst with ferric phosphate they bury themselves in the soil, making the farmers underestimate the efficacy of the product (Castle et al. 2017).

These strategies have contributed to a reduction in the number of metaldehyde failures in drinking water from 133 in 2014 to 4 in 2018 (Chief Inspector of Drinking Water 2019). Even though the total number of failures has been reduced in recent years, metaldehyde remains the single biggest challenge to the supply of pesticide-free water in the UK (Chief Inspector of Drinking Water 2018; Mohamad Ibrahim et al. 2019).

1.3.2 Chemical and physical methods for metaldehyde removal

Once metaldehyde enters an aquatic environment its degradation slows down and it becomes semi-persistent due to its physico-chemical properties (Bieri 2003) and is very difficult to effectively remove by current drinking water treatment options, including advanced treatment processes that require high energy and cost inputs (Cosgrove et al. 2019).

Xenobiotics that have a negative effect on drinking water quality are usually polar or semi-polar organic compounds at pg L^{-1} to µg L^{-1} concentrations (Benner et al. 2013). Metaldehyde is a polar molecule and its interactions with organic matter will be weak, becoming easily desorbed (Busquets et al. 2014). While classic strategies to remove contaminants such as the use of granular activated carbon (GAC) and powdered activated carbon (PAC) have been used for metaldehyde, their binding sites reach a saturation point relatively fast, and hence these processes are very costly to maintain at a large scale (Rolph et al. 2019).

Such constraints have pushed researchers and water companies to seek more advanced alternatives, such as specially designed materials such as nano-sized zinc composites or phenolic-based carbons (Busquets et al. 2014; Doria et al. 2013). However, most other removal methods combine UV advanced oxidation with another complementary approach such as UV/H2O2, UV/TiO2 and UV emitting diodes (Autin et al. 2012; Autin, Hart, et al. 2013a; Autin, Romelot, et al. 2013). However, all of these strategies are constrained by high energy demands (Castle et al. 2017).

1.3.3 Biotic degradation of metaldehyde

The biotic degradation of metaldehyde in treated soil or sediments proceeds at very different rates in different experiments; with DT_{50} values of between <1 and 67 d (Cranor 1990; Möllerfeld et al. 1993; X. Zhang and Dai 2006) according to some studies; while others have reported half-lives varying between 3.17–223 days depending on environmental conditions (Kay and Grayson 2014). These contrasting data are likely a result of the different methodologies used to measure the pesticide's persistence. Field studies usually result in lower half-lives than microcosm studies due to increased losses by runoff, erosion, and volatilization. Studies which use the compound applied in pellet form (as opposed to powder or a solution) would typically result in longer half-lives because of the time needed for the pellet to be weathered down and come into contact with the soil microorganisms (Bond 2018). The type of soil, pH, temperature,

and humidity can also influence the rate and extent of biological degradation and thus the recorded half-lives.

Degradation of metaldehyde occurs at very reduced rates in sterilized soils, which indicates that microbial activity is essential to this process (Simms et al. 2006). Furthermore it has been shown that in anaerobic soil incubations, metaldehyde is essentially stable (European Food Safety Authority 2010a). Moreover, an increased persistence of this compound has been observed in more anaerobic conditions such as those present in groundwater (Lapworth et al. 2012). These data indicate a predominantly biotic and aerobic microbial metabolism of the compound.

Balashova et al. (2020) used ^{14}C-radiorespirometry to assess the extent of metaldehyde mineralization across a range of agricultural, allotment and garden soil slurries, which ranged between 17.7 to 60.0% after 5 days. Their results suggested that catabolic competence to degrade metaldehyde was ubiquitous in the range of soils and specific conditions they tested and they hypothesized that the wide detection of metaldehyde in water sources might be more related with inadequate application regimes (runoff and fast leaching) rather than the inability of the microbial community to degrade the compound. Nonetheless a relatively high dose of metaldehyde was applied to each system (28 mg kg^{-1} soil) and the catabolic activity of the soil microorganisms under much lower concentrations should be tested.

Microbial activity is essential to metaldehyde degradation in the environment (Simms et al. 2006). The compound degrades to acetaldehyde, subsequently to acetic acid (acetate), which can be modified and used as a carbon and energy source by microorganisms, and ultimately to water and carbon dioxide, the usual final metabolites of aerobic degradation processes (Bieri 2003; Thomas 2016).

Biological sand filter experiments conducted by Rolph et al. (2019) indicated that an initial exposure to metaldehyde present in the water resulted in increased removal rates in subsequent exposures (acclimation). They showed that the removal was biological in nature and

their experiments suggested that only a small group of specialist degraders might be responsible for its uptake and removal in biofilters. These data offered a hypothesis as to why some biofilters can degrade metaldehyde for drinking water treatment whilst other do not.

Thomas et al. (2017) used domestic soils as starting material to isolate individual bacteria (*Acinetobacter* strain E1, *Variovorax* strain E3) capable of using metaldehyde as a sole source of carbon and energy (100 mg L^{-1}). *Acinetobacter* strain E1 was even capable of bringing down the initial concentrations to below the regulatory limit (0.1 µg L^{-1} metaldehyde) highlighting its potential to be used in future bioremediation strategies. Even though whole-genome sequencing was performed for both of these strains, the enzymatic basis for biological metaldehyde depolymerization could not be elucidated.

1.3.4 Evolution of genes for metaldehyde degradation

As mentioned before, the evolution of enzymes required for use of a compound as an energy source is one of the main drivers of bacterial adaptation to pesticide degradation. Bacteria develop new catabolic pathways in order to access new sources of nutrients, energy or to get rid of toxic products (Kolvenbach et al. 2014). Metaldehyde is a xenobiotic synthesized by man that did not exist previously in the environment. Despite its wider use as a pesticide only since the 1930's (Gimingham 1940), microbial degradation of metaldehyde exists (Balashova et al. 2020; Bieri 2003; European Food Safety Authority 2010b; Rolph et al. 2019; Thomas et al. 2017). Initially, one might think that microorganisms in the environment may have had relatively little time to evolve efficient mechanisms to degrade the compound. According to (Seffernick and Wackett 2001) initially, degradation rates of xenobiotics are slow and supported by non-specific reactions (cometabolism) which do not lead to microbial growth and may accumulate toxic by-products. Nevertheless, this does not seem to be the case anymore with metaldehyde since it has been shown that its degradation leads to complete aerobic mineralization products and it can support microbial growth on its own when used as a carbon source (European Food

Safety Authority 2010b; Thomas et al. 2017). These data indicate a more evolved pathway for the utilization of the compound has been achieved.

It has been noted that pesticides with high water solubility and low adsorptivity can usually develop higher degradation rates and extents of microbial growth more quickly, both conditions fulfilled by metaldehyde. Furthermore, structurally simple pesticides will develop degradation pathways in less time (Arbeli and Fuentes 2007).

Figure 1-2 summarizes the Innovation-Amplification-Divergence (IAD) model for enzyme evolution (Bergthorsson et al. 2007). Microbial enzymes that degrade new xenobiotics arise from promiscuous activities of previously existing enzymes with different roles in other pathways, that, even if they are initially not very efficient, they can substantially accelerate the reactions relative to the uncatalyzed state and provide a good starting point for further evolution to ultimately confer a selective advantage to the organism. Oftentimes the promiscuous activity results from the ability of the enzyme to bind and transform molecules that resemble the normal substrate (Copley 2009; O'Brien and Herschlag 1999). Even a few mutations can be capable of increasing the efficiency of an activity by several orders of magnitude (Copley 2020). Furthermore, the expression of the promiscuous enzymes should be able to occur under the specific environmental conditions in a given setting, which provides an advantage for constitutively expressed enzymes; and finally, the promiscuous enzyme must be able to continue to serve its original function. This oftentimes is achieved by gene duplication and specialization of one of the gene copies for the new substrate (Copley 2009). If paralogs become inefficient, then unnecessary gene copies are lost (Morgenthaler et al. 2019).

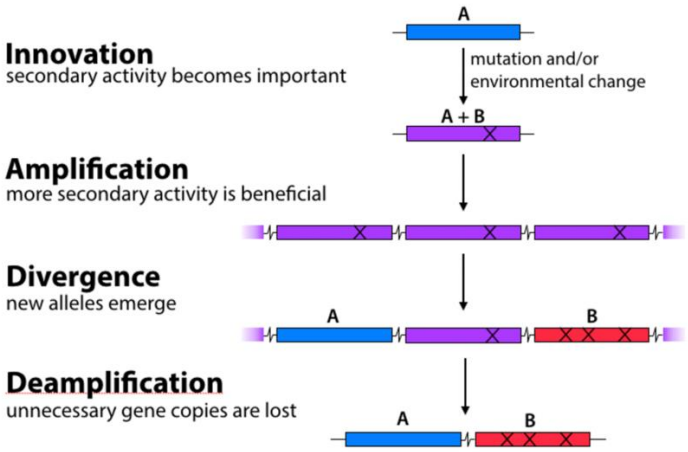


Figure 1-2. The Innovation-Amplification-Divergence (IAD) model of gene evolution. A mutation or an environmental change may generate that enzyme A acquires another physiologically relevant activity (B). The abundance of the modified enzyme is increased by gene amplification and further mutations can increase the efficiency of the new activity (B). If sufficient B activity is achieved, extra gene copies are usually lost. Taken from Morgenthaler et al. (2019).

Unfortunately, ancestral enzymes which were modified for xenobiotic degradation are rarely identifiable in nature. Nevertheless, two of the ones that have been identified are melamine deaminase (TriA), which gave rise to atrazine chlorohydrolase (AtzA), which share 98% identity (Seffernick et al. 2001), and dihydrocoumarin hydrolase, which evolved into methyl parathion hydrolase (G. Yang et al. 2019). However, this kind of identification is rarely possible given the enormous complexity and our still limited knowledge of microbial ecosystems (Copley 2020).

A degradative pathway with several reactions might evolve more quickly if all of these are harboured within a single microbe that can obtain the fitness advantage of the activity (Copley 2009) as it seems to be the case with metaldehyde (Thomas et al. 2017).

Given that horizontal gene transfer provides the main source of new enzymatic activities in prokaryotes and that it is easier for a metabolic ability to be shared between bacteria than it is to evolve multiple times (Ochman et al. 2000; Treangen and Rocha 2011), Thomas (2016) hypothesized that if several different metaldehyde-degrading isolates could be obtained, a straightforward way of identifying the metaldehyde-degrading genes would be to pinpoint

enzymes shared between the degrading isolates, but absent from closely related non-degrading strains. Unfortunately, the low number of degrading isolates obtained at the time and the higher cost of whole-genome sequencing hindered the ability to adequately test this hypothesis.

Once the enzymes responsible for metaldehyde degradation/mineralization are unequivocally identified, clues must be obtained to the function of the progenitor enzymes by sequence comparison or by testing their promiscuous activities. More importantly, degrading gene presence and expression could be used as an instrument to quantify metaldehyde degradation potential and activity in a given environmental setting, which would allow improved prediction of metaldehyde fate in the environment leading to better tailored policies and contingency plans in case of unexpected release events.

1.4 Detection and Quantification of Pesticide-degrading Genes in the Environment

The assessment and prediction of the fate of xenobiotics in the environment during bioremediation or natural attenuation can be carried out through different strategies: analysing the disappearance of the parent compound, detecting its transformation products or through evidence of transformation potential in a given setting (Fenner et al. 2013). It is in this last approach in which knowledge regarding degrading genes can be used to detect and quantify their copy numbers or their expression, which provides a measure of degradation activity. This objective could be accomplished once the identification of at least one set of metaldehyde-degrading genes is complete. The following section will focus on the strategies that could be used and the information that could be obtained via a combination of several different techniques once these genes are identified.

Several approaches for detection and quantification of specific xenobiotic degrading strains through marker (non-degrading) genes have been undertaken in the past (Widada et al. 2002). However, since pesticide degrading genes can be acquired by previously non-degrading

microbes (DiGiovanni et al. 1996) or lost from degrading organisms (Changey et al. 2011), detecting and quantifying the degrading genes themselves carries the obvious advantage of more accurately reflecting the status of the degrading community.

Different methods have been used for this purpose. Techniques involving DNA hybridization (using labelled DNA probes) were used in the past few decades to detect pesticide degrading genes in soil microbial populations (Holben et al. 1992; Parekh et al. 1995; Walia et al. 1990) ; however due to limitations mainly regarding their sensitivity and reliance on cultivation, they were gradually replaced by methods based on PCR amplification (Widada et al. 2002).

Initially, methods such as most probable number PCR (MPN-PCR) (Picard et al. 1992) were developed to provide a quantitative component to PCR amplifications and they have been used for assessment of contaminant-degrading genes (Salminen et al. 2008).

Improved quantitative PCR methods providing higher sensitivity and accuracy, such as quantitative competitive PCR (QC-PCR) and replicate limited dilution PCR (RLD-PCR) were also developed. QC-PCR relies on the inclusion of an internal control competitor to which each reaction is normalized (Piatak et al. 1993). The competitor molecules should be DNA fragments that share the same primer recognition sequences with the target yet yield distinguishable PCR products from the wild-type template (Zimmermann and Mannhalter 1996). This technique has been used to monitor the size of the atrazine-degrading bacterial community in soil, finding a transitory increase in the amount of *atzC* genes in soils pre-treated with atrazine (Martin-Laurent et al. 2003). However, the design of the competing molecule and the validation of the amplification efficiencies of the target and the competitor can be a laborious process (Heid et al. 1996). RLD-PCR relies on replicative dilution to extinction (Chandler 1998) and has been used to quantify hydrocarbon-degrading gene copies in soil (Tuomi et al. 2004), however it can be time and resource consuming.

1.4.1 Quantitative PCR (qPCR)-based Studies

Since its introduction and dissemination, most studies dealing with pesticide catabolic gene quantification have relied on real-time PCR also known as quantitative PCR (qPCR) (Klein 2002). qPCR allows the monitoring of the PCR amplification and quantification of its products in real time through fluorescence measurements. This technique can provide a sensitive, robust, and fast estimation of catabolic gene copy numbers in samples taken from pesticide bioremediation settings. Despite the obvious advantages of the method, the adoption and transfer of inadequate and varied protocols amongst researchers has led to problems when trying to reproduce data (Hayden et al. 2008; Jaworski et al. 2018; Taylor et al. 2019). Therefore, the minimum information for publication of quantitative real-time PCR experiments (MIQE) guidelines have been proposed to increase the uniformity and reliability of published research using this method (Bustin et al. 2009).

Numerous studies have used this technique to quantify copy numbers of pesticide-degrading genes to estimate the biodegradation potential and the response of the microbial community to these contaminants (Table 1-2). Besides pesticides, this method has also been used extensively for hydrocarbons; numerous studies have been published including protocol design and application in environmental settings for these compounds (J. Liu et al. 2019; Q. Liu et al. 2017; Meynet et al. 2015; Shahi et al. 2016; Shahsavari et al. 2016).

Table 1-2. Studies using qPCR to quantify pesticide-degrading gene copies in environmental samples

Pesticide Class	Main use	Pesticide	Gene	Enzyme(s)	Matrix	Complementary techniques	Main findings	Reference
Carbamates	Insecticides/nematicides	oxamyl, carbofuran	*cehA* *mcd*	carbamate hydrolase carbamate hydrolase	Agricultural soil	Parent compound quantification (HPLC/UV)	*ceh*A gene abundance and pH were negatively correlated with oxamyl half-life. Carbofuran stimulated *ceh*A and *mcd* gene abundance.	(Rousidou et al. 2017)
		oxamyl	*cehA*	carbamate hydrolase	Agricultural soil	Pesticide mineralization 16S rRNA gene amplicon sequencing 16S rRNA reverse transcription and amplicon sequencing	*ceh*A copy numbers increased concomitantly with oxamyl mineralization The diversity of the bacterial community was modified by oxamyl addition.	(Gallego et al. 2019)
Organochlorines	Insecticides/acaricides	hexachlorocyclohexane	*lin*A *lin*B	gamma-hexachlorocyclohexane dehydrochlorinase haloalkane dehalogenase	Dumpsite soil, agricultural soil	Parent compound quantification (GC-ECD) 16S rRNA gene T-RFLP	An indigenous community capable of HCH degradation was detected in the contaminated soils.	(Lal et al. 2015)
Organophosphates	Insecticides/nematicides/fungicides	chlorpyrifos	*mpd* *opd*	methyl parathion hydrolase parathion hydrolase	Agricultural soil	Parent compound quantification (GC-MS) Pesticide mineralization	Chlorpyrifos was degraded mainly through abiotic processes. Neither *mpd* nor *opd* sequences were detected.	(Ben Salem et al. 2018)
		tolclofos-methyl	*opd*	parathion hydrolase	Agricultural soil	Parent compound quantification (HPLC/UV)	*opd* gene copy number was significantly higher at week 4 than at week 0.	(Kwak et al. 2012)
Phenoxyalkanoic acids	Herbicides	2,4-D, MCPA, MCPP	*tfd*A	alpha-ketoglutarate-dependent dioxygenase	Arable soil, forest soil, other soil	Pesticide mineralization	Class III *tfd*A genes were most involved in mineralization of the pesticides	(Bælum and Jacobsen 2009)
		2,4-D, MCPA	*tfd*A	alpha-ketoglutarate-dependent dioxygenase	Agricultural soil, forest soil	*tfd*A gene diversity analysis	Unknown, diverse *tfd*A-like genes were abundant in soil samples.	(Zaprasis et al. 2010)
		MCPP	*tfd*A	alpha-ketoglutarate-dependent dioxygenase	Agricultural soil	Parent compound quantification (HPLC-UV)	MCPP half-life increased progressively with soil depth, with a considerable lag phase in subsoil.	(Rodríguez-Cruz et al. 2010)
		MCPA	*tfd*A	alpha-ketoglutarate-dependent dioxygenase	Urban soil, agricultural soil	Pesticide mineralization	A correlation was found between class III *tfdA* gene copy numbers and the rate of mineralization	(Nielsen et al. 2011)
		2,4-D, MCPA, MCPP	*tfd*A	alpha-ketoglutarate-dependent dioxygenase	Arable soil, forest soil, other soil	Pesticide mineralization	*tfd*A gene copy numbers involved in growth linked mineralization kinetics. Degradability of pesticides was 2,4-D > MCPA > MCPP	(Bælum et al. 2012)
Phenylureas/Sulphonylureas	Herbicides	chlorimuron-ethyl	*sul*E	sulfonylurea herbicide de-esterification esterase	Soil from abandoned land	Parent compound quantification (HPLC/UV)	An inoculated strain effectively contributed to pesticide removal.	(L. Yang et al. 2014)

						Quantification of N-cycling functional genes	Copy numbers of *sul*E genes decreased with time after inoculation.	
		diuron	*puh*A *puh*B	phenylurea hydrolase A phenylurea hydrolase B	River sediment Agricultural soil	Pesticide mineralization	Mineralization rates in soil were linked to *puh*B copies. *puh*A gene copies were not detected.	(Pesce et al. 2013)
		diuron	*puh*A *puh*B	phenylurea hydrolase A phenylurea hydrolase B	Agricultural soil with organic amendments	Parent compound quantification (HPLC/UV) A-RISA fingerprinting Enzymatic activity	Olive vermicompost increased enzymatic activities as well as bacterial and *puh*B gene abundance. Diuron pre-treatment correlated with high abundance of *puh*B gene copies.	(Castillo et al. 2016)
Triazines	Herbicides	atrazine	*atzA* *atzB* *atzC*	atrazine chlorohydrolase hydroxydechloroatrazine ethylaminohydrolase N-isopropylammelide isopropyl amidohydrolase	Agricultural soil	Pesticide mineralization RISA fingerprinting	Organic amendment did not modify atrazine mineralization and genetic potential. Structure of the bacterial community was significantly affected. Soils with similar genetic potential can exhibit different atrazine-degrading activities, which depended on soil.	(Martin-Laurent et al. 2004)
		atrazine	*atzA*	atrazine chlorohydrolase	Rhizosphere soil	Pesticide mineralization Parent compound quantification (GC/MS)	Addition of a bioaugmentation strain leads to more efficient mineralization of atrazine	(Thompson et al. 2010)
		atrazine	*trzN* *atzB* *atzC*	triazine hydrolase Hydroxydechloroatrazine ethylaminohydrolase N-isopropylammelide isopropyl amidohydrolase	Agricultural soil	Parent compound quantification (HPLC/UV)	Bioaugmentation could achieve a quick and significant reduction in atrazine concentrations	(Wang et al. 2013)
		atrazine	*trzN*	triazine hydrolase	Agricultural soil	Parent compound quantification (GC/FID)	Degrading gene quantification correlated with degradation efficiencies.	(Sagarkar et al. 2013)
		atrazine	*atzA* *atzB* *trzN*	atrazine chlorohydrolase hydroxydechloroatrazine ethylaminohydrolase triazine hydrolase	Boreal soil, agricultural soil	Pesticide mineralization	Low levels of atrazine degrading genes in boreal soils; no mineralization detected	(Nousiainen et al. 2014)
		atrazine	*trzN*	triazine hydrolase	Agricultural soil	Parent compound quantification (HPLC/UV) 16S rRNA gene amplicon sequencing	In most soils a single dose of the pesticide induced accelerated degradation. Neutralization of a low pH soil restored accelerated degradation.	(Yale et al. 2017)

The gene quantification approach is usually complemented with other strategies to maximize the information obtained from the system under study: either disappearance of the parent compound, detection of transformation products or both. Many studies also carry out microbial community analysis through fingerprinting or amplicon sequencing techniques to observe modifications in the microbial community structure of the soil.

Most of these studies have found a correlation between degrading gene copy number and pesticide degradation/mineralization activity (Bælum et al. 2012; Rousidou et al. 2017; Sagarkar et al. 2013), highlighting the value of this approach for biodegradation monitoring. Pre-exposed soils usually harboured higher degrading gene copies than non-exposed soils (Castillo et al. 2016; Lal et al. 2015). Very low or undetectable numbers of degrading genes have been observed in instances where elimination was negligible (Nousiainen et al. 2014), others in which degradation was mainly through abiotic processes (Ben Salem et al. 2018), or where degradation was presumably occurring through alternative biological pathways for pesticides such as diuron (Pesce et al. 2013). Notably, soils with similar genetic degrading potential can exhibit different pesticide-degrading activities, depending on factors such as soil type (Martin-Laurent et al. 2004) and pH (Yale et al. 2017). This emphasizes the value of assessing biodegradation potential/activity using several different strategies simultaneously.

Even though qPCR has long been the workhorse for quantification of functional genes in bioremediation, it is not without limitations. Most notably, DNA extracts from environmental samples often contain diverse contaminants and different backgrounds that can variably affect the activity of the Taq polymerase, producing misleading results (Carvalhais et al. 2013). This can sometimes be remediated by sample dilution, which will likely lead to undetectable target levels or by extensive sample purification, which is not always feasible (Rački et al. 2014).

1.4.2 Digital PCR (dPCR)-based Studies

Newer technologies with distinctive advantages have emerged. One of such technologies is digital PCR (dPCR), in which the sample is partitioned in droplets or distributed in nanolitre chambers and then a standard Taq polymerase PCR reaction is used to amplify target DNA. However, different to qPCR, data is acquired at the reaction endpoint (Taylor et al. 2017). This minimizes the influence of sample contaminants on the amplification and leads to more accurate and reproducible results when compared with qPCR, and allows for very low target quantitation (Rački et al. 2014; Taylor et al. 2015). dPCR is yet to be used for pesticide-degrading gene quantification, however it has been used for the determination of alkane hydrocarbon degrading genes (*alk*B1) in a pilot scale biopile field experiment at freeze-thaw temperatures (J. Kim et al. 2018). The authors performed nutrient amendment in one of the biopiles and found that degrading gene copy numbers were higher in the treated biopile versus the untreated one during the seasonal freezing and thawing phases. More studies using this technology for xenobiotic-degrading gene quantification will undoubtedly emerge in the coming years as it becomes increasingly available.

1.4.3 Loop Mediated Isothermal Amplification (LAMP)-based Studies

Loop mediated isothermal amplification (LAMP) has been used as an alternative to qPCR as a point-of-care, lower cost, diagnostic tool. LAMP is performed at constant temperature, so it removes the need for thermal cycling equipment, it is faster than PCR due to the high amplicon yield and can amplify target DNA even in complex matrices that can inhibit PCR (Notomi et al. 2000); detection can be visual or via fluorescence. Visual based LAMP was combined with most probable number for the quantification of reductive dehalogenase genes in groundwater samples (Kanitkar et al. 2017) using centrifuged cells instead of DNA extracts and a strong correlation was found with qPCR results.

1.4.4 Reverse Transcription Quantitative PCR (RT-qPCR)-based Studies

All the techniques discussed so far suffer from an important limitation: they reveal a genetic potential which is only expressed under favourable conditions. A more powerful way to monitor biodegradation involves the measurement of degrading gene expression, routinely carried out by reverse transcription-quantitative polymerase chain reaction (RT-qPCR). Here, RNA instead of DNA is used as a target for amplification. Since RNA cannot be used as a template for PCR directly, reverse transcription of the RNA template into cDNA is performed and then followed by the exponential amplification as done in qPCR (Bustin 2000). Studies using RT-qPCR to quantify expression of pesticide-degrading genes in environmental samples are shown in Table 1-3.

Several studies have found good correlations between degrading gene expression and pesticide mineralization activity (Monard et al. 2013; Nicolaisen et al. 2008). Moreover, while studying MCPA mineralization in agricultural soil samples, researchers found that *tfd*A mRNA levels were had a much better correlation with pesticide mineralization than *tfd*A gene copy numbers per gram of soil (Bælum et al. 2008). Even so, care should be taken in cases where the degrading gene is constitutively expressed, because the quantification of its expression would merely reflect the amount of active degraders, rather than active biodegradation (Monard et al. 2013). In these instances, if expression is normalized to gene copy numbers, this ratio has been shown to remain stable (Albers et al. 2015).

Table 1-3. Studies using RT-qPCR to quantify pesticide-degrading gene expression in environmental samples

Pesticide Class	Main use	Pesticide	Gene	Enzyme(s)	Matrix	Complementary techniques	Main findings	Reference
Benzonitriles	Herbicides	Dichlobenil	*bbd*A	Amidase	Sand filter material, groundwater	BAM quantification (HPLC-UV)	Bioaugmentation produced significant BAM degradation for 2-3 weeks.	(Albers et al. 2015)
						Total bacteria, *bbd*A gene copy and MSH1 strain quantification (qPCR)	Ratio of *bbd*A mRNA per *bbd*A gene (amidase expression) was stable over time.	
						Protozoan quantification (MPN)	Protozoa were observed to grow in the sand filters.	
Phenoxyalkanoic acids	Herbicides	2,4-D, MCPA	*tfd*A	Alpha-ketoglutarate-dependent dioxygenase	Agricultural soil	Pesticide mineralization	Relatively high degree of correlation was found between functional gene expression and mineralization rates, not so with gene copy numbers.	(Bælum et al. 2008)
						*tfd*A gene copy quantification (qPCR)		
						*tfd*A gene DGGE		
		MCPA	*tfd*A	Alpha-ketoglutarate-dependent dioxygenase	Agricultural soil	Pesticide mineralization	Transient maximums of gene expression were observed only during active mineralization.	(Nicolaisen et al. 2008)
						*tfd*A gene copy quantification (qPCR)	Microbial degrader activity correlated with *tfd*A mRNA presence.	
Triazines	Herbicides	Atrazine	*atzA*	atrazine chlorohydrolase	Agricultural soil	Pesticide mineralization	No *atz*A expression was detected in the tested soils.	(Monard et al. 2010)
			*atz*D	cyanuric acid amidohydrolase		16S rRNA gene copy quantification (qPCR)	Expression of *atz*D was greater in the soil with the highest atrazine mineralization activity	
		Atrazine	*atzD*	cyanuric acid amidohydrolase	Agricultural soil	Pesticide mineralization	Relative gene expression was positively correlated with the maximum rate of pesticide mineralization.	(Monard et al. 2013)
						16S rRNA gene copy quantification (qPCR)		

1.4.5 Metagenomic and Metatranscriptomic-based Studies

Even though the amplification-based approaches discussed so far greatly increase the sensitivity to allow the quantification of degrading genes in environmental samples, they may introduce specific biases and are subject to limitations. For instance, protein-coding genes in the environment usually have high variability in their nucleotide sequences (when compared with small subunit rRNA genes) (Gaby and Buckley 2017), and therefore, primers that target functional genes are sometimes designed to be degenerate in sequence to be able to encompass that variability (Jin and Mattes 2011; Morales et al. 2020). It has been shown that degenerate qPCR primers are subject to PCR bias (i.e. certain templates are favoured in their amplification and therefore their quantification will not reflect their true original template ratios), which can lead to dramatic mis-estimation of real gene copy numbers, even more than 10,000-fold in some instances (Gaby and Buckley 2017). Furthermore, amplicon-targeted methods are usually low throughput technologies, in which a small number of targets can be detected per run and *a priori* knowledge of the degrading genes is essential for primer design (Forbes et al. 2017). Newer methodologies have found a way to circumvent these issues.

The fact that DNA sequencing technologies are becoming increasingly affordable and accessible has led to a growing number of studies using metagenomic strategies to evaluate pesticide contaminated environments (Regar et al. 2019; Sangwan et al. 2014). Metagenomics involves the direct sequencing of extracted DNA from a sample (Gorski et al. 2019) and can be used to assess and quantify functional genes in environmental matrices without targeted amplification and without the use of sequence-specific primers. Essentially, it provides a detailed survey of all the genes that exist within a particular community. Nonetheless, some of its limitations include that it heavily samples the dominant microbes in the community and members with low abundance are sparsely covered (Techtmann and Hazen 2016) or not at all depending on sequencing power. Furthermore, data analysis can be very complicated and may involve genes with no homologs in bioinformatic databases (Delmont et al. 2012).

Metagenomic analysis has been used to evaluate the genetic pool present in pesticide contaminated environments and quantify its functional potential. In a study focusing on hexachlorocyclohexane, researchers characterized the metagenome of the soil microbial community of three sites with different contamination levels (Sangwan et al. 2012). Among different analyses, they selected twelve genes known to be involved in the degradation pathways for this pesticide and quantified the matching reads. The two sites with the highest contamination levels had a higher metabolic potential to degrade hexachlorocyclohexane and higher relative abundance of degrading genes (*lin* genes) when compared with the least polluted one. Another study analysing activated sludge samples from pesticide wastewater treatment plants (Fang et al. 2018) detected 68 subtypes of pesticide degrading genes, which together showed a relative abundance from 2.08 to 7.14%, which varied depending on the seasons and pesticide wastewater properties. They also found that activated sludge from pesticide wastewater treatment plants had a higher abundance of pesticide degrading genes compared to others in which saline or freshwater was treated.

Researchers have also evaluated the metagenome of agricultural soils with a history of chlorpyrifos exposure, nevertheless at different times and intensity (Jeffries et al. 2018), while also undertaking chlorpyrifos degradation assays in the laboratory. Relative abundance of metabolic pathways resulted in clustering of samples in two distinct groups which coincided with effective and ineffective control of pests in the field and also with slow and rapid experimental degradation of the pesticide in the samples. Soils with faster degradation harboured a distinct community with increased nutrient cycling and transport pathways, as well as enzymes putatively involved in the metabolism of phosphorus.

Another methodology that has been gaining increasing momentum is metatranscriptomic analysis. It has the advantage over metagenomics that it indicates the specific microbial genes being expressed at the moment of sampling. It involves RNA extraction from an environmental sample, conversion into cDNA (akin to RT-qPCR) and sequencing in a similar manner to

metagenomics (Techtmann and Hazen 2016). Metatranscriptomic analysis has been used to assess gene transcripts linked with the degradation of aromatic compounds and pesticides from wheat rhizosphere sample sequences obtained from the EBI metagenomics database (D. Singh et al. 2018). The researchers found abundant transcripts associated with degradation of xenobiotics (including pesticides) aromatic amines, carbazoles, benzoates, naphthalene, ketoadipate pathway, phenols and biphenyls.

Newer, more informative and accurate methods are being applied by researchers to understand the role of microorganisms in polluted environmental systems. Lower costs in sequencing and advances in molecular methods have allowed scientists to gain insight into xenobiotic-degrading genes and pathways, not only related to culturable microorganisms but increasingly to genes associated to the yet to be cultured microbes.

However, the most informative genetic and biochemical analyses can still only be performed on culturable organisms. To be able to obtain the most information out of environmental systems, efforts must be made to effectively isolate and culture these environmentally relevant microbes. This would allow a great amount of information to be added to genomic and proteomic bioinformatic databases which are, ultimately, the ones used in metagenomic studies to assign gene functions.

Furthermore, monitoring studies that aim to assess not only the genetic potential, but also microbial taxonomic shifts, parent compound and metabolite concentrations, physical and chemical parameters of the environment have and will continue to prove most informative, especially if done in the field, and should be encouraged. Once the enzymes involved in metaldehyde degradation are identified, these strategies can be applied to obtain the most information out of environmental systems.

1.5 Bioaugmentation of slow sand filters for metaldehyde removal

Once enough information is gathered regarding metaldehyde degrading strains and their respective catabolic genes, especially tailored biological solutions for drinking water purification, alone or combined with other methods, could be designed, and put into practice.

Typical large-scale drinking water treatment steps include screening, aeration, coagulation and flocculation, sedimentation, filtration, and disinfection. It is in the (sand) filtration step in which biological strategies seem more promising for micropollutant purification.

The effectiveness of the sand filtration step in water purification is associated with physicochemical and microbial processes (Bai et al. 2013). As water is filtered through the sand, both biotic and abiotic materials accumulate in between the spaces of the sand grains and a biofilm eventually forms (Rooklidge et al. 2005). Inorganic species and residual particles are removed or transformed and the biofilm can play an important role in the transformation or degradation of harmful compounds (Bai et al. 2013; Cakmakci et al. 2008; Clasen 1998).

Two types of sand filtration are used: rapid sand filters and slow sand filters. In the first, loading rates are of approximately 1-10 mh^{-1} and 0.1-1 mh^{-1} in the latter (Fitzpatrick and Gregory 2003) which correspond to hydraulic retention times of minutes and hours respectively.

The autochthonous microbiota of sand filters on their own can transform a wide range of xenobiotics with varying efficiencies, from no removal to nearly complete removal as it has been tested at different scales (Arvin et al. 2004; Corfitzen et al. 2009; Ho et al. 2007; Rolph et al. 2019; Zearley and Summers 2012). However longer contact times, usually associated with slow sand filters can generate enhanced removal of the problematic compound (Zearley and Summers 2012; Zuehlke et al. 2007). The main problem with this strategy on its own lies in the fact that microbial communities might only develop the ability of efficiently degrading a specific compound, if at all, after long periods of exposure to the xenobiotic, and might generate

transformation products that may be more toxic or persistent than the original compound (Benner et al. 2013; Helbling et al. 2010; Prasse et al. 2011; Rolph et al. 2019).

Given these limitations, bioaugmentation has emerged as a viable alternative. Bioaugmentation comprises the addition of microorganisms to enhance a specific biological activity (Vogel 1996). Moreover, the use of bioaugmentation strategies with specific pesticide degrading strains or communities has gathered the interest of researchers to introduce missing catabolic genotypes, to accelerate pesticide degradation activity, and to avoid the production of toxic metabolites (Dejonghe et al. 2000; Sniegowski et al. 2011).

Bioaugmentation of slow sand filters with specialized degrading strains or populations emerges as a promising strategy to deal with micropollutants in drinking water sources. Haig et al. (2014) showed that the microbial community and water quality performance of laboratory scale slow sand filters can replicate that of full-scale filters. Also, bioaugmentation of laboratory scale and larger slow sand filters and similar systems for the degradation of other xenobiotics has been successful in the past for compounds such as microcystin, geosmin, 2-methylisoborneol and estrogen (Bourne et al. 2006; Haig et al. 2016; McDowall et al. 2009).

If this strategy is pursued for any compound it must obviously be preceded by the isolation and characterization of individual strains or consortia that are capable of degrading the problematic pollutant. It would be optimal if the bioaugmentation agents are capable of complete mineralization of the compound, to avoid dealing with problematic transformation products which can be more toxic and or persistent than the original compound. Also, most candidate strains for bioaugmentation procedures have only been shown to utilize their specific substrate at relatively high concentrations (much higher than micropollutant levels) (Benner et al. 2013). Thomas et al. (2017) already showed that metaldehyde-degrading organisms with these two important characteristics (mineralization capability and low concentration degradation capacity) are present in nature. Moreover, obtaining a wider repertoire of degrading strains with versatile capabilities and testing them in much more challenging

conditions is also an important objective of this thesis, which will allow the selection of optimal strains for bioaugmentation purposes.

Pre-culturing of the strain(s) to obtain optimal microbial loads in the sand filters would be necessary, but the success of the strategy would depend on the ability of the bioaugmentation agent to integrate into the complex microbial community of the filter, attach to the filter media, survive predation and competition and metabolize the pollutant even when challenged with fluctuating concentrations (Benner et al. 2013). Such complex set of conditions and interactions so far can only be tested experimentally but its likelihood of positive outcomes can be optimized by choosing the most promising strains.

1.6 Aims of the project

Valuable information has been obtained in recent investigations regarding the use of microorganisms as a potential biotechnological solution to the problem of metaldehyde contamination in drinking water. However, a deeper understanding of the microorganisms involved in the process, the metabolic pathways that they use to degrade the compound, the potential transfer of these traits between organisms, and the adaptation of the processes to a larger scale are challenges that are still ahead. To address these issues, the following aims have been proposed:

- To isolate a broad collection of metaldehyde-degrading strains from soil and, through comparative genomics, identify potentially horizontally-transferred metaldehyde-degrading enzymes (Chapter 2).
- To detect changes in the prokaryotic populations in soil as a response to metaldehyde applications (Chapter 2).

- To determine the genetic context of the identified metaldehyde-degrading genes, including whether they are chromosomally located or are part of mobile genetic elements and gain insight into their probable evolutionary history (Chapter 3).
- To develop a qPCR analytical technique to be able to quantify the copy number of metaldehyde-degrading genes in soil and relate it to metaldehyde removal (Chapter 3).
- To test the metaldehyde-degrading strain collection to determine if they are capable of lowering metaldehyde concentrations below the regulatory limit in batch assays with raw water and sand (Chapter 4).
- To use the best performing strains for pilot scale slow sand filter assays to determine their ability to remove metaldehyde from a continuous flow of raw water to levels below regulatory limits (Chapter 4).
- To include a final discussion focusing on the accomplishments of the project and future perspectives regarding biological metaldehyde degradation (Chapter 5).

CHAPTER 2: GENOMIC BASIS FOR METALDEHYDE DEGRADATION REVEALED BY SELECTION, ISOLATION AND CHARACTERISATION OF A LIBRARY OF DEGRADING STRAINS FROM SOIL[2]

2.1 Introduction

Soils contain extremely diverse microbial communities with versatile metabolic capabilities. The acquisition of new traits occurs by diversification of the existing genetic material from the metagenome and is further enabled by horizontal gene transfer (Maheshwari et al. 2017). Thus, it is possible for novel metabolic activities to emerge and be evolutionarily reinforced via selection within soil microbial communities (Fierer 2017; Kuzyakov and Blagodatskaya 2015). Such processes are important to agriculture as pesticides are used to improve crop productivity, but also impact soil microbes, which may degrade them. Genes whose products degrade novel chemicals and confer a competitive nutritional advantage on their microbial hosts have potential to be selected for in soils (Arbeli and Fuentes 2007).

Metaldehyde $(CH_3CHO)_4$ is a pesticide used worldwide as a molluscicide to control snails and slugs that damage agricultural crops and domestic gardens (Eckert et al. 2012). It is a xenobiotic that has been used as a pesticide since the 1930s (Gimingham 1940); it is hydrolytically and photolytically stable (Carpenter 1989; Kegley et al. 2016). Metaldehyde is applied to many crops, including oilseed rape, wheat and potatoes, and accounted for 86% of all molluscicide applications in the UK by area in 2016 (Garthwaite et al. 2018). It is applied as a pelleted bran bait in the autumn when molluscs prosper due to the humid conditions. Rainfall can dissolve metaldehyde and carry the compound to watercourses which may be used for drinking water abstraction (Lazartigues et al. 2013).

[2] A modified version of Chapter 2 has been published as a peer-reviewed journal paper:
Castro-Gutiérrez, V., Fuller, E., Thomas, J.C., Sinclair, C.J., Johnson, S., Helgason, T., Moir, J.W.B., 2020. Genomic basis for pesticide degradation revealed by selection, isolation and characterisation of a library of metaldehyde-degrading strains from soil. Soil Biology and Biochemistry 142. doi:10.1016/j.soilbio.2019.107702.
The published version of the manuscript has been included as Annex 1.

Metaldehyde is recalcitrant to removal using conventional drinking water treatment processes based on adsorption of substances to activated carbon (Castle et al. 2017). Moreover, detections of metaldehyde have been the main cause for already treated water not meeting pesticide standards since its monitoring began (Chief Inspector of Drinking Water 2017b). Hence, water companies have been exploring alternative solutions for metaldehyde detection and elimination. Therefore, the idea of using biological strategies to detect and degrade metaldehyde has recently emerged. Metaldehyde can be quickly degraded in soils (Lewis et al. 2016; H. Zhang et al. 2011) and is oxidised to carbon dioxide under aerobic conditions (European Food Safety Authority 2010a), in comparison to its long half-life in sterile conditions (Simms et al. 2006). This suggests a strong involvement of microorganisms in its degradation. The biotic degradation of metaldehyde in treated soil or sediments proceeds at very different rates in different experiments; with DT_{50} values of between <1 to 67 d (Cranor 1990; Möllerfeld et al. 1993; X. Zhang and Dai 2006).

Very few studies have investigated further the mechanism underpinning the microbial degradation of metaldehyde. Rolph (2016) found improved metaldehyde removal in slow sand filter material following a period of acclimation with relatively elevated metaldehyde concentrations. Thomas and collaborators (2017) were the first to isolate bacterial strains capable of metaldehyde degradation (*Acinetobacter calcoaceticus* E1 and *Variovorax* E3). Even so, the diversity of metaldehyde degraders found in nature, the biochemical mechanisms responsible for its degradation, and the effect metaldehyde inputs have in shaping microbial communities in biological habitats remain largely unexplored. In this chapter the diversity of organisms capable of degrading metaldehyde using culture-based and culture-independent molecular methods is examined, focusing on the impact of historical and experimental application of the pesticide. A set of genes, shared amongst several independent isolates, responsible for metaldehyde degradation is identified.

2.2 Materials and Methods

2.2.1 Chemicals and reagents

Metaldehyde (99%) was purchased from Acros Organics, NJ; sodium sulfate (≥99%) was purchased from Honeywell-Fluka, Bucharest; all other chemicals were purchased from Sigma-Aldrich, St. Louis, MO.

2.2.2 Analytical methods for metaldehyde

2.2.2.1 Metaldehyde extraction and quantification in aqueous matrices

Aqueous samples were centrifuged (4000 g, 10 min) and 0.4 mL supernatant added to 0.5 mL dichloromethane in glass chromatography vials (Thermo Scientific), vortexed for 30 s and stabilized for 20 min. 5 µL of organic phase were injected into an Agilent 7820A gas chromatograph (Stockport, UK) fitted with a HP-5 column and Flame Ionization Detector. Chromatographic parameters were described previously (Tao and Fletcher 2013). Limit of detection and limit of quantification were 0.05 mg L^{-1} and 0.15 mg L^{-1} respectively. Calibration curves were constructed with standard metaldehyde solutions in minimal salts media (MSM) (Thomas et al. 2017) and sample peak area was interpolated in the calibration curves.

2.2.2.2 Metaldehyde extraction and quantification in soil

Each sample for extraction consisted of 10 g of soil in glass Falcon tubes (Kimble-Chase, NJ). 5 g anhydrous sodium sulfate were added, contents homogenized, and 15 mL dichloromethane added. Tubes were vortexed for 30 s and placed in ultrasonic bath for 20 min. Centrifugation (2000 rpm, 10 min) followed. 2 mL of supernatant were withdrawn from each system, passed through sodium sulfate/silanized glass wool column, and filtered through a 0.45 µm PTFE filter. Chromatographic analysis was as described for aqueous samples. Peak area for samples was interpolated in a linear calibration curve. Limit of detection and limit of quantification were 0.25 mg kg^{-1} soil and 0.75 mg kg^{-1} soil respectively.

2.2.3 Soil sampling and characterization

Three soil sample collections were performed. First collection, January 2017: two council-run allotment plots (Hob Moor Allotments, York, UK; 53.946113, -1.103654). Plot J-Met had been regularly treated with metaldehyde for at least 5 years. Plot J-NoMet had not been treated with metaldehyde for at least 5 years. Second collection, June 2017: a lettuce and cilantro-producing agricultural farm (Caballo Blanco, Cartago, Costa Rica; 9.854884, -83.901134) periodically exposed to metaldehyde for at least 3 years (C-Met). Third collection, October 2017: two different allotment sites: Scarcroft Allotments (S) (York, UK; 53.950840, -1.092036) and Hob Moor Allotments (H). For each site, two strawberry plots were chosen: one treated with metaldehyde during the July-August period (S-Met and H-Met) and another plot with no metaldehyde applications for at least 5 years (S-NoMet, H-NoMet). In each plot, triplicate 300 g sub-samples were taken from the top 10 cm of the soil and stored in plastic bags. These were combined into a composite sample and stored in loosely tied plastic bags in a temperature-controlled room at 23°C until analysed. Soil samples were air dried overnight and sieved (2 mm mesh). Physical and chemical parameters for soils are shown in Table 2-1.

Table 2-1. Physical and chemical characteristics of allotment soils used in this chapter

Plot	Silt (%)	Clay (%)	Sand (%)	Soil Class	Organic Matter (%)	pH	Max WHC (% dry mass)
J-Met	48.5	12.1	39.4	Sandy Silt Loam	3.5	6.7	-
J-NoMet	42.0	10.3	47.7	Sandy Silt Loam	4.0	6.5	-
H-Met	46.1	10.6	43.3	Sandy Silt Loam	5.6	6.3	70.9
H-NoMet	40.3	8.8	50.9	Sandy Loam	4.9	6.9	85.9
S-Met	56.3	10.4	33.3	Sandy Silt Loam	7.4	7.2	83.3
S-NoMet	44.9	7.8	47.3	Sandy Silt Loam	8.5	7.2	116.9

2.2.4 Microbiological analyses

Figure 2-1 shows an overview of methods used for the microbiological analyses of the collected soil samples. Soils J-Met, J-NoMet and C-Met were used for isolation of metaldehyde-degrading bacterial strains only. Soils S-Met, S-NoMet, H-Met, H-NoMet were additionally used

for constructing soil metaldehyde degradation profiles and analysing associated changes in the microbial communities.

2.2.4.1 Selective enrichment for the isolation of metaldehyde-degrading strains (a)

A modified selective-enrichment procedure (Abraham and Silambarasan 2013) was performed to obtain metaldehyde-degrading bacterial strains from the soils. For the first enrichment passage (P1), 1 g of soil was inoculated in 100 mL of MSM with added metaldehyde as sole source of carbon at 100 mg L^{-1} in 250 mL Erlenmeyer flasks. All metaldehyde-containing media was filter sterilized (0.22 µm) prior to use and supplemented with 2.0 mL L^{-1} trace elements (Vishniac and Santer 1957). P1 was incubated with orbital shaking (25°C, 150 rpm) in the dark for 72 h. Subsequently, 1.0 mL was transferred to 100 mL fresh medium (P2) and incubated as described for P1. Third and fourth passages (P3 and P4) were carried out similarly. Aliquots from P3 and P4 were streaked onto solid supplemented MSM (0.75% agarose) with 75 mg L^{-1} metaldehyde in triplicate and incubated for 72 h at 25°C in the dark. Different colony morphotypes that grew on this media were purified by restreaking. Pure cultures were also streaked on MSM plates with no added carbon source to identify possible agarose degraders and oligotrophs, which were discarded after their inability to degrade metaldehyde was corroborated in liquid media. Strains that grew on metaldehyde plates and not in no added carbon plates were presumptively identified as metaldehyde degraders. Isolates were preserved by freezing at -80°C in 15% glycerol.

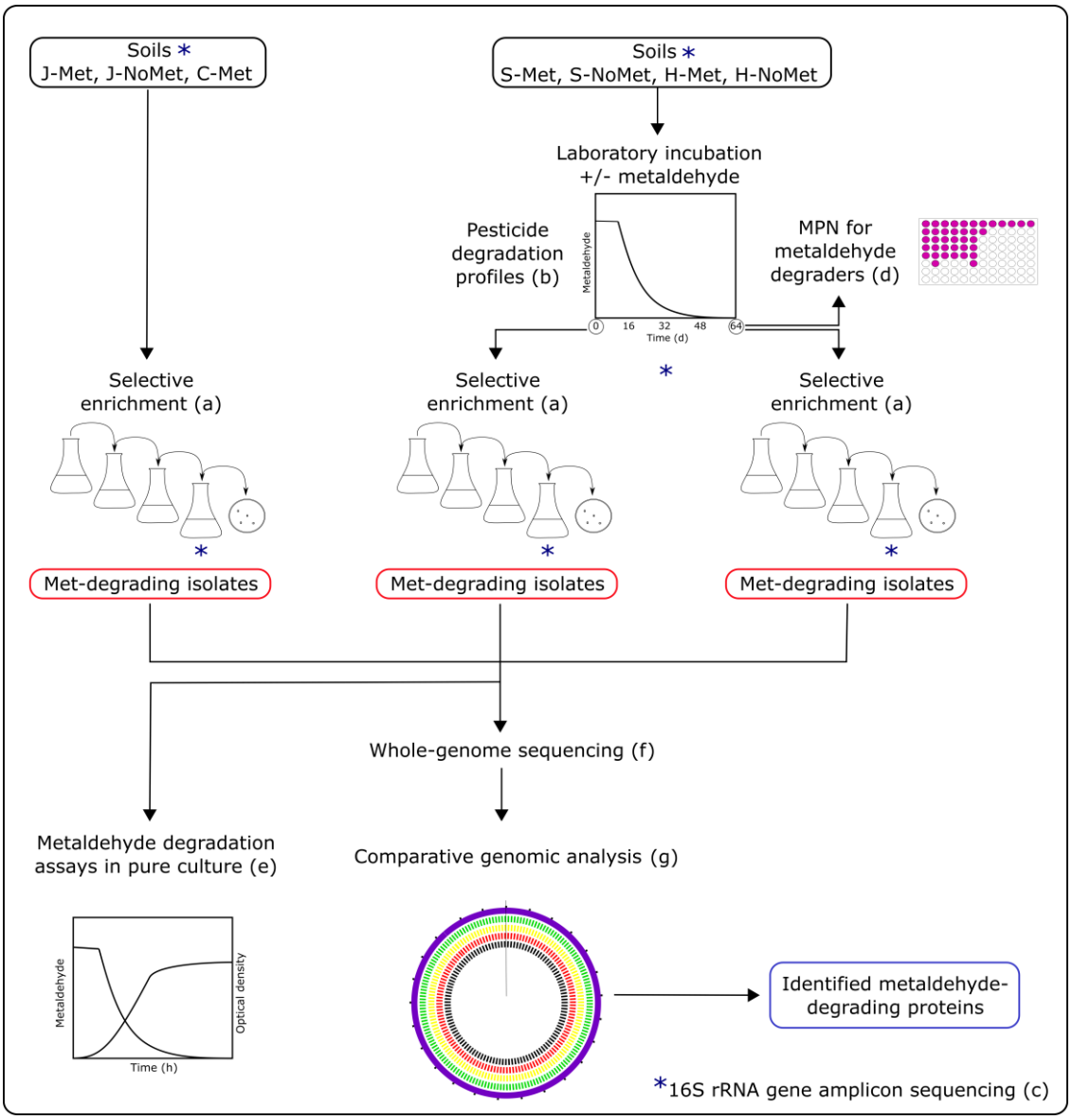


Figure 2-1. Overview of microbiological analyses performed for soil samples. All soils were subjected to a selective enrichment process in minimal medium with metaldehyde as the only source of carbon for the isolation of metaldehyde-degrading strains (**a**). Specific soils were also incubated in the laboratory with and without an initial addition of metaldehyde; remaining pesticide concentrations were periodically quantified for the construction of soil metaldehyde degradation profiles (**b**). 16S rRNA gene amplicon sequencing analyses (**c**) were performed for the allotment soils before, during and after the laboratory incubations with and without metaldehyde to evaluate changes in the microbial communities as a consequence of metaldehyde exposure. After incubation, the number of metaldehyde-degrading microorganisms was determined using a most-probable number technique (**d**). Soils subjected to laboratory incubation were also used for selective enrichment to isolate metaldehyde-degrading strains. Metaldehyde-degrading capabilities of the isolated strains were tested through metaldehyde degradation assays in pure culture (**e**). Whole-genome sequencing was performed for the metaldehyde-degrading strains (**f**); comparative genomic analysis (**g**) was carried out for the identification of candidate genes involved in metaldehyde-degradation.

2.2.4.2 Metaldehyde degradation profiles and microbial community changes in allotment soil microcosms (b)

Metaldehyde degradation was followed in allotment soil samples S-Met, S-NoMet, H-Met, H-NoMet after single and repeated metaldehyde applications at 15 mg kg^{-1} soil to investigate accelerated degradation of metaldehyde. Samples were taken for soil genomic DNA extraction to determine changes in soil microbial community in response to metaldehyde addition and incubation of sieved soils. Controls with no metaldehyde addition were labelled cS-Met, cS-NoMet, cH-Met and cH-NoMet. The specific protocol is subsequently described in detail.

Metaldehyde degradation after a single application and associated microbial community changes

Metaldehyde degradation was followed in freshly collected allotment soil samples S-Met, S-NoMet, H-Met, H-NoMet after a single initial metaldehyde application. Before the start of the assay, soil moisture was adjusted to 40% of the MWHC and soil was stored in the dark, in loosely tied plastic bags for two weeks for the microbial community to stabilize.

Three different sets of samples per soil were prepared. One set was treated with metaldehyde and used for pesticide quantification analyses. A second set was also treated with metaldehyde but used for genomic DNA extraction. A third set, not treated with metaldehyde, was used as a control for DNA analyses. For each set and time point, 10 g of soil were added to triplicate 50 mL glass Falcon tubes. A 1.5 mg mL^{-1} solution of metaldehyde in methanol was prepared and 100 µL were added to soil tubes used for pesticide analyses and for genomic DNA extraction after pesticide exposure. Nominal starting pesticide concentration was 15 mg kg^{-1} soil. Controls in which only methanol was added were named cS-Met, cS-NoMet, cH-Met and cH-NoMet. Methanol was allowed to evaporate from all systems for 3.5 h. Soil in each tube was then manually homogenized and the mass for all systems was recorded. Incubation was carried out in the dark at 25±2°C and 100% relative humidity for 64 days. At time points 0, 2, 4, 8, 16,

32 and 64 d triplicate systems were stored at -20°C until metaldehyde was quantified. At time points 0, 16, 32 and 64 d, single samples for DNA analyses and their respective controls for each soil were stored at -20°C until DNA extraction was performed. Constant soil moisture was maintained by additions of sterile deionized water when necessary as determined by mass measurements.

Metaldehyde quantification was carried out as described previously to obtain metaldehyde degradation profiles. Data regression for curve fit was performed using CAKE (Computer Assisted Kinetic Evaluation, Syngenta, UK) version 3.3.6. Regression model was chosen according to FOCUS work group guidelines (FOCUS 2006). Single First Order and Modified Hockey stick models were chosen where appropriate.

Accelerated degradation of metaldehyde and associated microbial community changes

Degradation assays to test for accelerated biodegradation of metaldehyde were carried out in a similar manner as detailed above but using soil samples after a 6-month storage (25°C) and with an additional metaldehyde application of 15 mg kg^{-1} soil 64 d after the start of the assay. Metaldehyde was additionally quantified at times 0, 1, 2, 4, 8, and 16 d after the second metaldehyde application for the S-Met soil and at times 0, 2, 4, 8, 16 and 32 d after the second application for the rest of the soils (S-NoMet, H-Met, H-NoMet). At time points 0, 64 and 96 d, triplicate samples for DNA analyses and their respective controls for each of two soils (S-Met and S-NoMet) were stored at -20°C until DNA extraction was performed.

2.2.4.3 16S rRNA gene amplicon sequencing analyses (c)

Genomic DNA was extracted from the original soils, from soils during metaldehyde degradation assays and from passage P4 of all selective enrichments for community 16S rRNA gene amplicon sequence analyses, as well as from the pure cultures of metaldehyde-degrading strains for identification.

Genomic DNA was extracted from soils J-Met, and C-Met and from the end of stage P4 of all 100 mg L^{-1} enrichment cultures using Powersoil DNA isolation kit (Qiagen) following the manufacturer's instructions. Genomic DNA from soils S-Met, S-NoMet, H-Met and H-NoMet was extracted using NucleoSpin Soil DNA extraction kit (Macherey-Nagel, Düren, Germany). For these soils, a relative 16S rRNA gene copy number quantification technique (Smets et al. 2016) was carried out by spiking them before extraction with an internal DNA standard (*Thermus thermophilus* DSMZ 46338) at an estimated 1% of total DNA. Genomic DNA from presumptive metaldehyde-degrading strains was obtained using DNeasy Blood and Tissue kit (Qiagen, Venlo, Netherlands) according to the manufacturer's instructions. The quality of nucleic acid extractions was verified in SybrSafe-stained (Thermo Fisher Scientific) 1.0% agarose gels. DNA concentration and purity were quantified using a NanoDrop spectrophotometer (Thermo Fisher Scientific).

To identify the presumptive metaldehyde-degrading strains 16S rRNA genes were amplified as described elsewhere (Castro-Gutiérrez et al. 2016) and amplicons were sent for Sanger sequencing at Source Bioscience (Nottingham, UK) or GATC Biotech (Constance, Germany) on an ABI 3730xl DNA Analyzer. Partial 16S rRNA gene sequences for presumptive metaldehyde-degrading strains were assembled into contigs and edited with DNA Baser v4 (Heracle BioSoft). Phylogenetic affiliations were determined by BLASTN searches (Z. Zhang et al. 2000) against strains in the prokaryotic type strain 16S ribosomal RNA gene sequence database of NCBI and named as a given species based on greater than 98.65% identity to the closest match (M. Kim et al. 2014).

Analysis of whole community 16S rRNA amplicon sequences

For whole community 16S rRNA amplicon analyses the V4-V5 region of the 16S rRNA gene was amplified for each soil and each enrichment culture using 515F/806R primers (Walters et al. 2016) PCR conditions and purification steps were described elsewhere (Castro-Gutiérrez

et al. 2018) 16S rRNA amplicons were sequenced on an Illumina MiSeq instrument using 300-bp paired-end sequencing at the Bioscience Technology Facility of the University of York.

Whole-community 16S rRNA V4-V5 region amplicon sequence information for the soils and all enrichment cultures was quality checked using Fastqc software v. 0.10.1. Read sequences were analysed using QIIME2 v2017.12 (Caporaso et al. 2010). Demultiplexed paired-end sequences were imported and the Divisive Amplicon Denoising Algorithm 2 was implemented for quality filtering (Q score ≥ 30) , chimera removal, and feature table construction (Callahan et al. 2016). Taxonomy was assigned to the feature table using the Greengenes 13.8 reference database (McDonald et al. 2012). Feature tables were rarefied to equal number of reads, and the associated taxonomy was then extracted for further analysis.

PRIMER7 (Primer-E Ltd., Auckland, New Zealand) was used for all statistical analyses. Abundance data was standardized by samples, transformed by fourth root and a Bray-Curtis similarity matrix was constructed. Principal coordinates analysis (PCO) was used for data ordination. Permutational MANOVA (PERMANOVA) was used assess the influence of different factors in community composition (9999 permutations). PERMDISP was used to test for homogeneity of multivariate dispersions.

2.2.4.4 Enumeration of metaldehyde-degrading microorganisms in allotment soils (d)

The number of aerobic culturable metaldehyde-degrading microorganisms in allotment soil samples after laboratory incubation +/- metaldehyde was determined using a most probable number (MPN) enumeration technique in microtitre plates (Dinamarca et al. 2007). Allotment soils S-Met, S-NoMet, H-Met and H-NoMet exposed to metaldehyde during metaldehyde degradation assays and their respective no metaldehyde controls (cS-Met, cS-NoMet, cH-Met and cH-NoMet) were analysed.

2.2.4.5 Metaldehyde degradation assays in pure culture (e)

To confirm the ability of isolated strains to degrade metaldehyde, bacterial growth and metaldehyde degradation assays were performed in MSM with 150 mg L^{-1} metaldehyde. Inocula were prepared by growing each strain on nutrient agar for 72 h and resuspending growth in MSM to initial optical density at 600 nm (OD_{600nm}) = 0.1. 0.625 mL of this inoculum were added to 100 mL metaldehyde-supplemented MSM in triplicate 250 mL Erlenmeyer flasks. Abiotic controls with an equal volume of MSM instead of inoculum were also prepared. Cultures were incubated in an orbital shaker at 25°C and 150 rpm in the dark. Samples (1 mL) were withdrawn from the triplicate independent cultures at different time points for OD_{600nm} measurement and metaldehyde quantification.

2.2.4.6 Whole-genome sequencing of metaldehyde-degrading strains (f)

Confirmed metaldehyde-degrading strains were sent for whole-genome sequencing at MicrobesNG (Birmingham, UK). Sequencing was performed on Illumina MiSeq platform using 2x250bp paired-end reads. *De novo* assembly and quality assessment were performed using SPAdes (Bankevich et al. 2012) and QUAST (Gurevich et al. 2013) respectively. Automated fast annotation was performed using Prokka (Seemann 2014).

2.2.4.7 Identification of candidate genes involved in metaldehyde-degradation pathways (g)

To identify candidates for proteins involved in the metaldehyde-degradation pathway, the annotated genomes from newly isolated metaldehyde-degrading strains (*Acinetobacter bohemicus* JMET-C, *Acinetobacter lwoffii* SMET-C, *Pseudomonas vancouverensis* SMET-B, *Caballeronia jiangsuensis* SNO-D), that of the previously identified degrader *A. calcoaceticus* strain E1 (Thomas et al. 2017) and closely related non metaldehyde-degrading reference strains (*A. calcoaceticus* RUH2202 and *A. bohemicus* ANC3994) were compared. Inability of these latter

reference strains to degrade metaldehyde was corroborated experimentally. Reference strains were purchased from the Leibniz Institute DSMZ culture collection. A previously developed Python script (JC Thomas; available at https://pypi.org/project/blast-score-ratio/) was used to identify proteins shared between degrading strains but absent from the non-degrading strains. Using this tool, BLAST score ratio (BSR) (Rasko et al. 2005) was calculated for each of the annotated proteins of *A. calcoaceticus* E1 against the most similar proteins present in each of the other degrading and non-degrading *Acinetobacter* strains, *P. vancouverensis* SMET-B, and *C. jiangsuensis* SNO-D. Proteins with a BSR value equal to 0.9 or more (shared between the degrading strains) and with a BSR value lower than 0.45 in the non-degrading strains were chosen as candidate proteins and listed. The results of the proteome comparison analyses were corroborated, and figures generated using the BSR-based PATRIC Proteome Comparison Service (Wattam et al. 2017).

2.2.5 Accession numbers

Raw reads for 16S rRNA amplicons and draft whole-genome sequencing data for metaldehyde-degrading strains were deposited in the European Nucleotide Archive under study PRJEB30540. Specific sequences for metaldehyde-degrading genes are available in Annex 2.

2.3 Results

2.3.1 Initial enrichment cultures and sporadic isolation of metaldehyde-degrading strains

An overview of the experimental approach used in this work is shown in Figure 2-1. A selective enrichment procedure in liquid media with metaldehyde as sole source of carbon was applied to soil samples from three separate soil collections to isolate metaldehyde-degrading strains. Changes in the abundance of taxa generated by this process were assessed by 16S rRNA gene amplicon sequence analysis (Table 2-2).

Table 2-2. Most abundant taxa present at the end of the last passage (P4) of the successful selective enrichment cultures with metaldehyde as the sole source of carbon (100 mg L^{-1}) and their respective abundance in the original soil used as inoculum

Soil	Family	Genus	Abundance in soil (%)	Abundance in enrichment (%)
J-Met†	Pseudomonadaceae	*Pseudomonas*	0.5	21.1
	Moraxellaceae	*Acinetobacter**	ND	20.3
	Weeksellaceae	*Chryseobacterium*	ND	19.0
	Flavobacteriaceae	*Flavobacterium*	0.3	9.1
	Comamonadaceae	Unassigned	0.5	8.0
	Comamonadaceae	*Hydrogenophaga*	ND	7.2
C-Met	Comamonadaceae	*Acidovorax*	0.2	24.5
	Sphingomonadaceae	*Sphingobium**	ND	21.8
	Burkholderiaceae	*Burkholderia*	0.1	17.6
	Flavobacteriaceae	*Flavobacterium*	ND	15.5
	Pseudomonadaceae	Unassigned	ND	5.2
	Chitinophagaceae	*Niabella*	ND	3.3
H-Met	Sphingomonadaceae	*Sphingobium**	0.0	35.3
	Nocardiaceae	*Rhodococcus**	0.1	21.5
	Comamonadaceae	*Comamonas*	ND	9.1
	Oxalobacteraceae	*Cupriavidus*	0.0	6.3
	Alcaligenaceae	*Achromobacter*	ND	5.6
	Xanthomonadaceae	*Stenotrophomonas*	0.1	5.4
H-NoMet	Oxalobacteraceae	Unassigned	ND	50.8
	Nocardiaceae	*Rhodococcus**	0.1	40.5
	Rhizobiaceae	Unassigned	0.1	2.6
	Caulobacteraceae	*Brevundimonas*	ND	2.4
	Alcaligenaceae	*Achromobacter*	0.0	1.7
	Micrococcaceae	Unassigned	0.9	0.4
S-Met	Moraxellaceae	*Acinetobacter**	ND	94.1
	Pseudomonadaceae	*Pseudomonas**	0.3	3.5
	Burkholderiaceae	*Pandoraea*	ND	0.7
	Xanthomonadaceae	*Lysobacter*	0.1	0.2
	Clostridiaceae	*Clostridium*	0.6	0.2
	Alcaligenaceae	*Achromobacter*	ND	0.2
S-NoMet	Burkholderiaceae	*Burkholderia**	ND	75.5
	Comamonadaceae	*Hydrogenophaga*	ND	10.0
	Rhizobiaceae	Unassigned	0.1	4.8
	Pseudomonadaceae	*Pseudomonas*	0.4	3.7
	Pseudomonadaceae	Unassigned	0.1	1.4
	Hyphomicrobiaceae	*Devosia*	0.2	1.0

ND: not detected

* A member of the taxon was subsequently isolated from the enrichment and confirmed to be a metaldehyde degrader

† Isolates of the genera *Pseudomonas*, *Comamonas*, *Rhodococcus* and *Leucobacter* were recovered from passage P4 of this enrichment and were shown to be unable to use metaldehyde as a sole source of carbon in pure liquid cultures.

The complete list of metaldehyde-degrading strains successfully isolated by the enrichment culture procedure throughout the whole study and their respective identification is

presented in Table 2-3. Not surprisingly, the taxa corresponding to the subsequently isolated metaldehyde-degrading strains tended to dominate the composition of the final stage of the enrichment cultures. However, in some cases other taxa made up a considerable proportion of the community. These may constitute metaldehyde-degrading strains that cannot be isolated in solid media using this approach or non-degrading strains that are using metaldehyde degradation products or other by-products of degrading strains as sources of carbon (Neilson and Allard 2012).

Table 2-3. Metaldehyde-degrading strains isolated by enrichment culture procedure and identification based on 16S rRNA gene sequences against the NCBI database (limited to sequences from type material).

Soil of origin	Strain code	Closest relative	Closest relative GenBank Accession No.	Similarity (%)	No. of bases compared	Identification of the isolate
J-Met	JMET-C	*Acinetobacter bohemicus* ANC 3994(T)	KB849175	99.9	1406	*Acinetobacter bohemicus*
C-Met	CMET-H	*Sphingobium chlorophenolicum* NBRC 16172	NR_113840.1	98.5	1329	*Sphingobium* sp.
H-Met	HMET-A	*Rhodococcus globerulus* NBRC 14531(T)	BCWX01000023	100.0	1366	*Rhodococcus globerulus*
H-Met	HMET-G	*Sphingobium chlorophenolicum* NBRC 16172	NR_113840.1	98.4	1356	*Sphingobium* sp.
H-NoMet	HNO-A	*Rhodococcus globerulus* NBRC 14531(T)	BCWX01000023	100.0	1370	*Rhodococcus globerulus*
S-Met	SMET-B	*Pseudomonas vancouverensis* ATCC 700688(T)	AJ011507	98.9	1387	*Pseudomonas vancouverensis*
S-Met	SMET-C	*Acinetobacter lwoffii* NCTC 5866(T)	AIEL01000120	98.7	1394	*Acinetobacter lwoffii*
S-NoMet	SNO-D	*Burkholderia jiangsuensis* MP-1[1]	NR_133991.1	99.9	1398	*Caballeronia jiangsuensis*

[1] The taxon *Burkholderia jiangsuensis* has been reclassified as *Caballeronia jiangsuensis* (Dobritsa and Samadpour 2016).

For soils from first collection, isolation of a metaldehyde-degrading strain (*A. bohemicus* strain JMET-C) was successful from metaldehyde-exposed soil (J-Met) (Table 2-3), while no metaldehyde-degrading strains were isolated from the non-exposed soil (J-NoMet). For the second soil collection, performed in a previously-exposed agricultural soil from Costa Rica (soil C-Met), metaldehyde-degrading strain *Sphingobium* sp. strain CMET-H was successfully isolated. For allotment soils from the third collection, which included metaldehyde exposed and non-exposed soils (H-Met, S-Met, H-NoMet, S-NoMet), no metaldehyde-degrading strains were initially isolated. Thus, at that point no degraders could be isolated from non metaldehyde-exposed soils and the isolation of metaldehyde degraders from soils previously exposed to metaldehyde was sporadic.

2.3.2 Metaldehyde is degraded faster in soils after metaldehyde treatment

Metaldehyde degradation profiles for allotment soils H-Met, H-NoMet, S-Met and S-NoMet after a single application are shown in Figure 2-2a. To quantify the persistence of metaldehyde in the soil samples, data regression for the degradation profiles was performed using Single First Order or modified Hockey-stick models (Table 2-4) (FOCUS 2006). With a half-life of 3.9 d, degradation was much faster in soil S-Met than other soils.

Figure 2-2b shows metaldehyde degradation profiles for allotment soils subjected to two consecutive pesticide applications after a 6-month storage. A single pesticide application was enough to generate 1.7 to 6.1-fold reductions in metaldehyde half-lives in all the allotment soil samples on the second application; however, the effect was more evident in the soils in which the half-life was initially higher. The results show accelerated degradation occurred in the samples, which highlights the importance of biological mechanisms for metaldehyde elimination in soil.

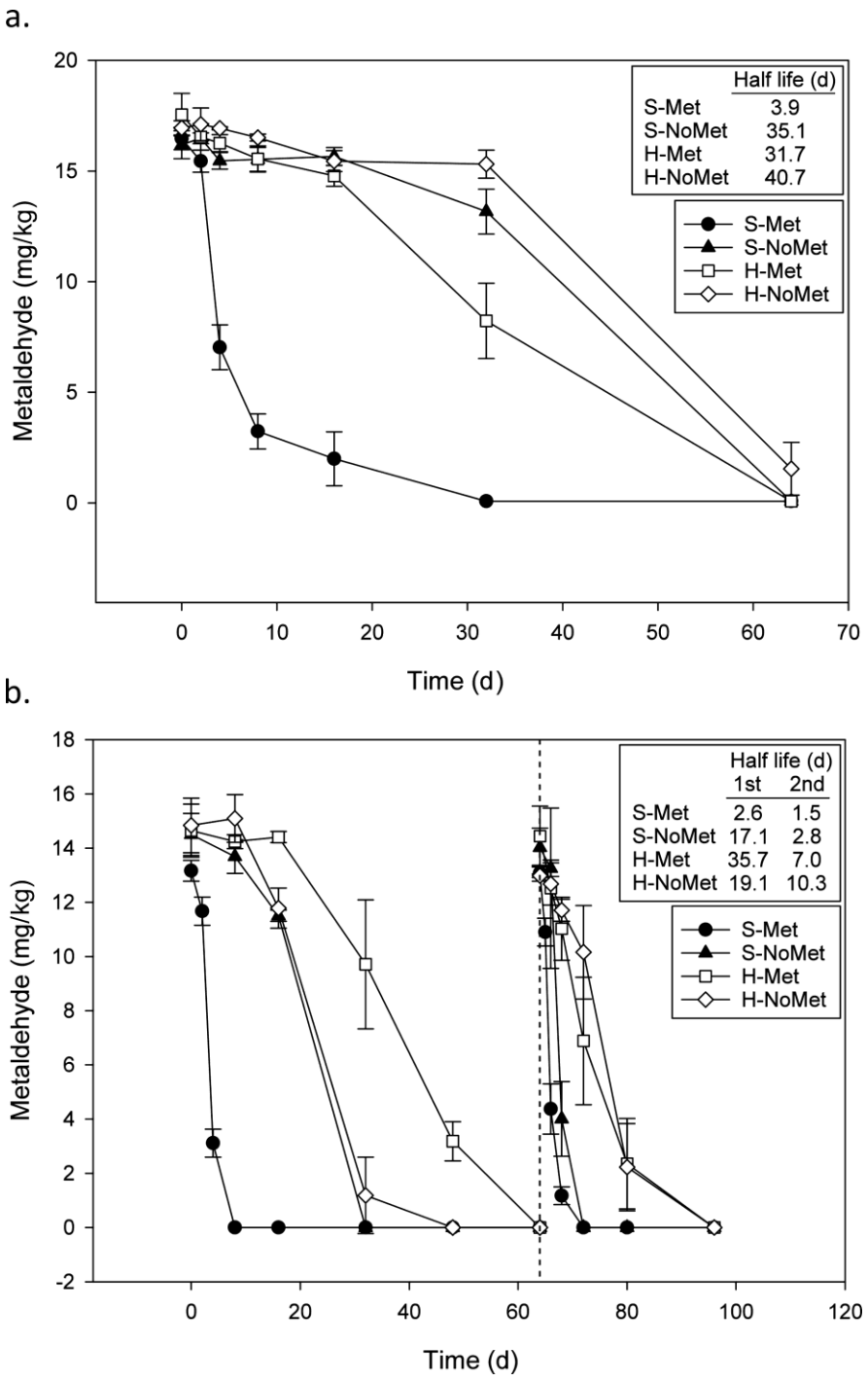


Figure 2-2. a. Metaldehyde degradation profiles in freshly collected allotment soils after an initial application of 15 mg kg^{-1} soil. **b.** Degradation of metaldehyde allotment soils after storage (6 months) following an initial application of 15 mg kg^{-1} soil and a second application (dashed vertical line) at the same dose. Bars represent standard deviation for 3 replicates. Regression statistics are presented in Table 2-4. S-Met: Scarcroft Metaldehyde, S-NoMet: Scarcroft No Metaldehyde, H-Met: Hob Moor Metaldehyde, H-NoMet: Hob Moor No Metaldehyde.

Table 2-4. Regression statistics for Single First Order and modified Hockey-stick fittings of metaldehyde removal in laboratory-incubated allotment soil samples.

Soil	Storage	Metaldehyde application	Regression model	k_1 (d^{-1})	k_2 (d^{-1})	Half-life (d)	r^2
S-Met	Fresh	Single	SFO*	-	0.1776	3.9	0.9373
S-NoMet	Fresh	Single	MHS†	0.002523	0.1615	35.1	0.9974
H-Met	Fresh	Single	MHS†	0.010070	0.1468	31.7	0.9984
H-NoMet	Fresh	Single	MHS†	0.003852	0.0735	40.7	0.9976
S-Met	6 months	First	SFO*	-	0.2711	2.6	0.9067
S-NoMet	6 months	First	MHS†	0.007238	0.4058	17.1	0.9999
H-Met	6 months	First	MHS†	0.001042	0.0793	35.7	0.9963
H-NoMet	6 months	First	MHS†	8.4×10^{-12}	0.1447	19.1	0.9998
S-Met	6 months	Second	SFO*	-	0.4745	1.5	0.9548
S-NoMet	6 months	Second	SFO*	-	0.2517	2.8	0.8953
H-Met	6 months	Second	SFO*	-	0.0998	7.0	0.9893
H-NoMet	6 months	Second	MHS†	0.025290	0.1905	10.3	0.9995

* Single First Order
† Modified Hockey-stick

2.3.3 Consistent enrichment and isolation of a greater diversity of metaldehyde degraders after incubating soil microcosms with metaldehyde in the laboratory

The accelerated degradation of metaldehyde pointed towards evolutionary selection of biologically-driven degradation of metaldehyde. Hence, it was hypothesized that metaldehyde-degrading organisms have been further enriched in soils after exposure to metaldehyde in the laboratory. Thus, to improve the recovery of metaldehyde-degrading isolates, selective enrichment for metaldehyde degraders was performed using the soil samples already incubated with metaldehyde for 64 d. This strategy permitted the isolation of metaldehyde-degrading strains from all four allotment soils in which the selective enrichment procedure had initially failed. Strains from diverse genera, including Gram-negative (*Acinetobacter*, *Pseudomonas*, *Sphingobium*, *Caballeronia*) and Gram-positive (*Rhodococcus*) isolates were successfully obtained. The identified metaldehyde-degrading strains are listed in Table 2-3. For all soils, in total, six Gram-negative and two Gram-positive strains were isolated. Whole-genome

sequencing was performed for all the distinct taxa. Quality statistics for sequencing runs and assemblies are shown in Table 2-5.

Table 2-5. Quality statistics for draft whole-genome sequencing runs and assemblies of metaldehyde-degrading strains (based on contigs of size >= 500 bp).

Identification	Strain code	Median insert size (bp)	Mean coverage (fold)	Number of reads	# contigs	Largest contig (bp)	Total length (bp)	GC (%)	N50 (bp)	L50 (bp)
Acinetobacter bohemicus	JMET-C	344	185.2	1 583 338	76	377 650	3 673 201	39.5	142 826	9
Sphingobium sp.	CMET-H	457	30.9	442 607	164	278 173	5 626 381	63.6	75 454	23
Rhodococcus globerulus	HNO-A	585	21.8	376 674	100	416 713	6 905 491	61.7	161 159	15
Pseudomonas vancouverensis	SMET-B	361	185.4	2 966 426	102	474 267	6 417 132	59.8	133 520	14
Acinetobacter lwoffii	SMET-C	301	149.3	1 369 861	79	399 021	3 554 352	40.2	136 921	9
Caballeronia jiangsuensis	SNO-D	493	37.0	804 434	124	428 341	8 659 758	62.7	162 268	18

The most abundant taxa in the final stage of these successful enrichment cultures, and their respective percentages in these original soils are also shown in Table 2-2. In the original soils, the genera with the highest abundances that could be confidently assigned taxonomy across all samples included the archaea *Candidatus* Nitrososphaera (9.6% of the initial community on average, SD: 3.1%) *Bacillus* (4.3%, SD: 2.5%) and *Kaistobacter* (2.4%, SD:1.7%), all which are commonly reported as some of the main genera found in soil (Zhalnina et al. 2013) and whose numbers decreased to undetectable levels at the end of the enrichment cultures. On the other hand, even though some genera such as *Acinetobacter*, *Sphingobium* and *Burkholderia*, all well-known xenobiotic degraders, were undetectable in the original soils with the sequencing depth used, their numbers increased to be amongst the dominant members of the enrichment cultures, which highlights both the strong selection pressure the populations were subjected to, and their ability to respond rapidly to it (Kurm et al. 2017).

2.3.4 Variation in the bacterial community due to metaldehyde addition and laboratory incubation

The addition of metaldehyde to the soil samples and the incubation in-soil increased the chances of isolating metaldehyde degraders. To further explore the variations in the microbial community occurring because of these procedures, changes were assessed by two different techniques: Most-probable number (MPN) of metaldehyde degraders and 16S rRNA amplicon sequencing of DNA from soil.

Figure 2-3a shows the results of the MPN assay of metaldehyde-degrading bacteria at the end of the degradation assay. For all the allotment soils tested metaldehyde degraders were significantly more abundant on exposure to metaldehyde compared to untreated controls. For all the soils except S-Met (which showed the fastest metaldehyde degradation even without metaldehyde exposure) this was a difference of several orders of magnitude.

The abundances of 16S rRNA gene copies were compared between soil samples during the metaldehyde degradation time course (Figure 2-3b). Quality statistics for sequencing runs are shown in Annex 3. Irrespective of the addition of metaldehyde, there is a consistent marked increase in bacterial abundance throughout the incubation period for all soils, which may also increase the chances of successful isolation of metaldehyde degraders. The fraction of the population at the phylum level corresponding to Proteobacteria showed an increase for all soils between the 0 d and 16 d time points and showed a slow decline thereafter.

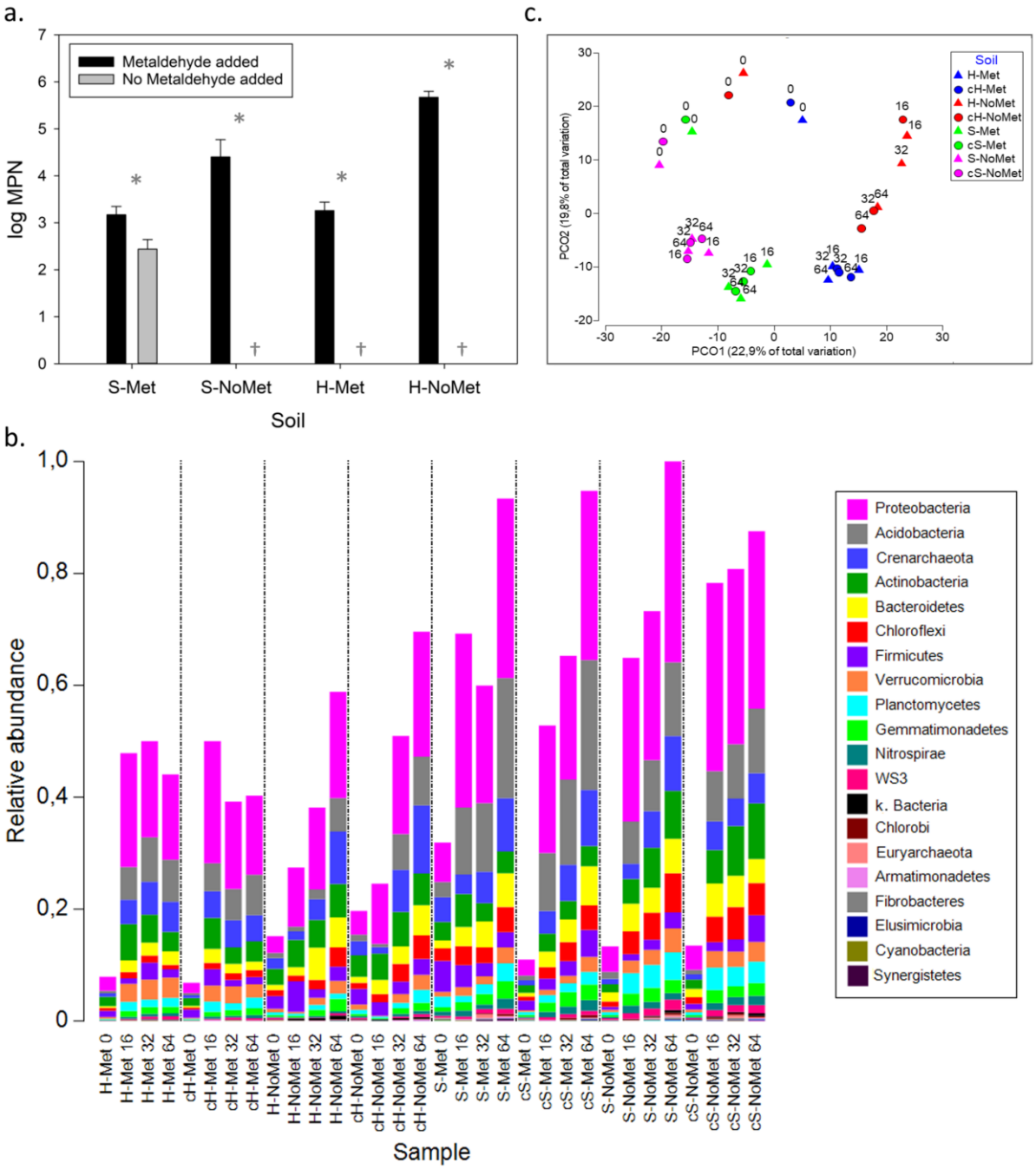


Figure 2-3. a. Most probable number of culturable metaldehyde-degrading bacteria per g of soil 64 d after the addition of metaldehyde at 15 mg kg^{-1} or no addition at all. Dagger indicates <23 MPN g^{-1} soil. Bars represent standard deviation for 3 replicates. Asterisk indicates significant difference ($p < 0.05$). **b.** Abundance and classification by phyla of 16S rRNA gene copies in individual allotment soil samples (relative to the sample with the highest count) with and without an initial addition of metaldehyde (15 mg kg^{-1}) throughout a 64 d incubation period. Only the most abundant 20 phyla are listed in the key. **c.** Principal coordinates analysis plot of whole bacterial community 16S rRNA gene amplicons from individual allotment soil samples with and without an initial addition of metaldehyde (15 mg kg^{-1}) throughout a 64 d incubation period. Labels indicate day of sampling. Samples were rarefied to 56749 sequences each after denoising and quality control.

To explore the differences in the composition of bacterial communities during the in-soil metaldehyde degradation assay, a PCO ordination of whole bacterial community 16S rRNA gene amplicons was constructed (Figure 2-3c). Principal coordinates 1 and 2 accounted for 42.7% of the variation. PERMANOVA analyses revealed that soil of origin (Pseudo-F = 6.26, p = 0.0001, df = 31) and time (Pseudo-F = 2.6058, p = 0.0001, df = 31) but not metaldehyde addition in the laboratory (Pseudo-F = 0.26951, p = 0.9999, df = 31) significantly influenced the soil community composition. Sample group dispersions were homogeneous for all factors as revealed by PERMDISP analyses.

Table 2-6 shows top ten taxa for each allotment soil for which abundance increases in samples exposed to metaldehyde in the laboratory versus non-exposed controls in the time frame from 16 to 64 d. The list is dominated by anaerobic organisms for most soils.

2.3.5 Microbial community changes in allotment soils during repeated applications

Figure 2-4a shows the relative abundance of 16S rRNA gene copies between the soil samples at different times and with repeated application. Irrespective of the addition of metaldehyde, bacteria and archaea were undergoing an increase in abundance throughout the incubation of the soils. A similar behaviour was observed during previous assays with single metaldehyde applications. Proteobacteria was the most abundant phylum for all soils and time points. Actinobacteria decreased from an average of 14.6 % of the community at day 0 to 6.3 % at day 96 for soil S-Met, while for soil S-NoMet it decreased from 11.7 % to 7.4 %. Bacteroidetes increased from an average of 5.4 % at day 0 to 9.3 % at day 96 for soil S-Met, while for soil S-NoMet it increased from 5.7 % to 13.7 %.

Changes in microbial populations during two consecutive applications of metaldehyde after a 6-month storage were visualised as a PCO ordination for soils S-Met and S-NoMet and the results are presented in Figure 2-4b. A good representation was obtained since principal coordinates 1 and 2 accounted for 67.0 % of the variation.

Table 2-6. Identification of top ten taxa for each allotment soil for which abundance increases in samples exposed to metaldehyde in the laboratory versus non-exposed controls in the time frame from 16 to 64 d.

Soil	Family	Genus	Species	Fold increase
H-Met	Tissierellaceae	*Tepidimicrobium*	Unassigned	1065
	Unassigned*	Unassigned	Unassigned	592
	Tissierellaceae	*Tissierella/Soehngenia*	Unassigned	342
	Sphingomonadaceae	*Sphingobium*	Unassigned	323
	Caldicoprobacteraceae	*Caldicoprobacter*	Unassigned	306
	Tissierellaceae	Unassigned	Unassigned	233
	Enterobacteriaceae	*Escherichia*	*Coli*	209
	Entotheonellaceae	Unassigned	Unassigned	190
	Clostridiaceae	*Clostridium*	*subterminale*	182
	OM60	Unassigned	Unassigned	105
H-NoMet	Paenibacillaceae	*Brevibacillus*	*reuszeri*	211
	Desulfovibrionacea	*Desulfovibrio*	Unassigned	82g
	Methanomassiliicoccaceae	*vadinCA11*	Unassigned	75
	Clostridiaceae	*Alkaliphilus*	Unassigned	58
	Oxalobacteraceae	Janthinobacterium	*lividum*	57
	Ectothiorhodospiraceae	Unassigned	Unassigned	52
	Nostocaceae	Unassigned	Unassigned	46
	Unassigned**	Unassigned	Unassigned	32
	Planococcaceae	*Sporosarcina*	Unassigned	24
	Thermomonosporaceae	*Actinoallomurus*	*iriomotensis*	24
S-Met	Methanomassiliicoccaceae	*vadinCA11*	Unassigned	403
	Desulfovibrionaceae	*Desulfovibrio*	Unassigned	395
	Clostridiaceae	*Natronincola/Anaerovirgula*	Unassigned	143
	Dethiosulfovibrionaceae	Unassigned	Unassigned	109
	Dethiosulfovibrionaceae	*Aminobacterium*	Unassigned	84
	Veillonellaceae	*Sporomusa*	Unassigned	84
	Patulibacteraceae	Unassigned	Unassigned	64
	Veillonellaceae	*vadinHB04*	Unassigned	57
	JTB215	Unassigned	Unassigned	55
	Eubacteriaceae	Unassigned	Unassigned	51
S-NoMet	Streptomycetaceae	*Streptomyces*	Unassigned	179
	Paenibacillaceae	*Paenibacillus*	*amylolyticus*	118
	Sphingomonadaceae	*Sphingomonas*	*wittichii*	116
	Pseudonocardiaceae	Unassigned	Unassigned	113
	Streptomycetaceae	Unassigned	Unassigned	104
	Clostridiaceae	*Alkaliphilus*	Unassigned	98
	Legionellaceae	*Legionella*	Unassigned	90
	Cytophagaceae	*Cytophaga*	Unassigned	80
	Clostridiaceae	*SMB53*	Unassigned	56
	Peptococcaceae	*Pelotomaculum*	Unassigned	52

* This taxon could only be identified to the order level: Thermoanaerobacterales

**This taxon could only be identified to the order level: 258ds10

Samples from H-Met, H-NoMet, S-Met, and S-NoMet soils were rarefied to 56749, 75440, 72945 and 90379 reads respectively.

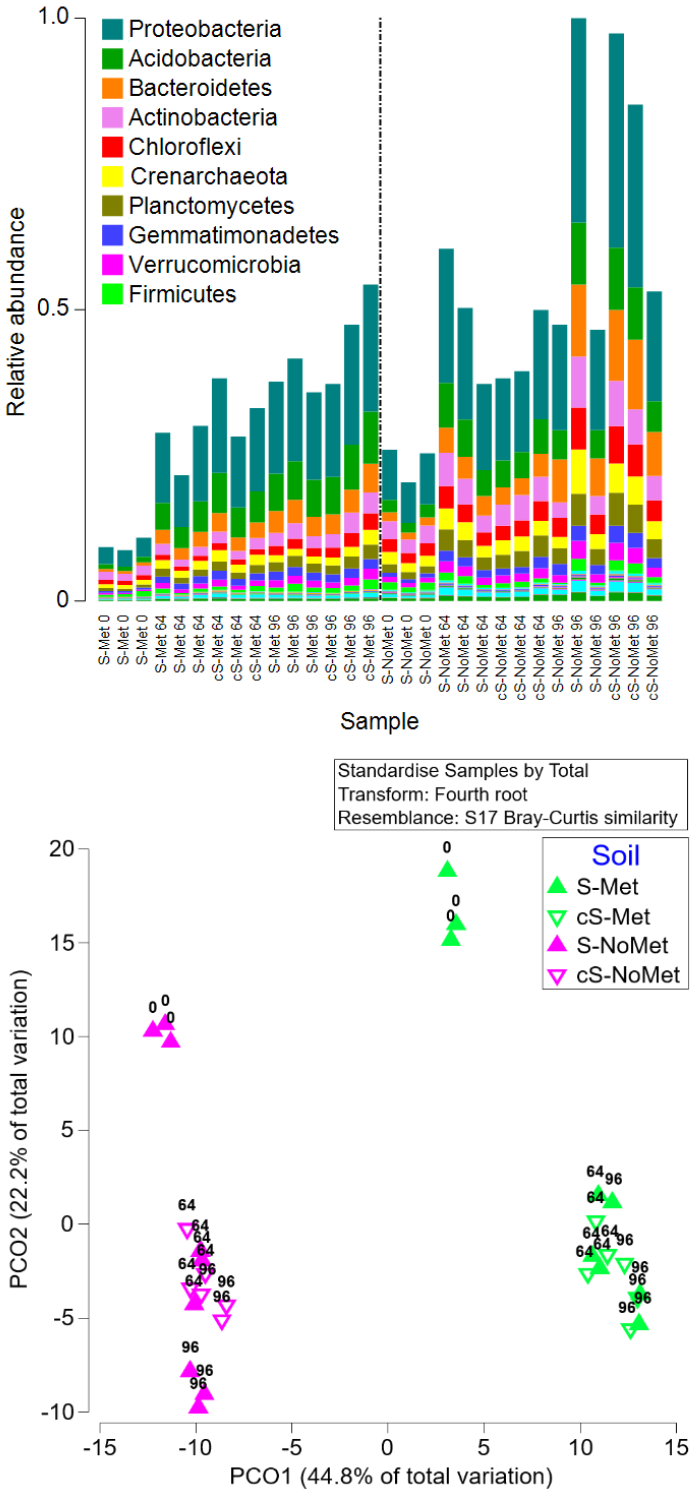


Figure 2-4. a. Abundance of 16S rRNA gene copies in soil samples (relative to the sample with the highest count) with additions of metaldehyde (15 mg kg^{-1}) at times 0 d and 64 d (S-Met and S-NoMet) and their respective controls without metaldehyde addition (cS-Met and cS-NoMet). Different colours indicate different phyla (restricted to the top 10 taxa). **b.** Principal coordinates analysis plot of whole bacterial community 16S rRNA gene amplicons from allotment soil samples with additions of metaldehyde (15 mg kg^{-1}) at times 0 d and 64 (S-Met and S-NoMet) and their respective controls without metaldehyde (cS-Met and cS-NoMet). Labels indicate day of sampling. Samples were rarefied to 64225 sequences each after denoising and quality control.

PERMANOVA analyses could not detect an effect of repeated metaldehyde addition on the differences in the overall composition of the microbial community (Pseudo-F = 1.0452, p = 0.341, df = 29). Restricting the analysis to phylum Proteobacteria only did not change this conclusion (Pseudo-F 1.3625, p = 0.224, df = 29). For the factors time and soil of origin, PERMANOVA analyses were significant (Pseudo-F = 4.4152, p = 0.001, df = 29 and Pseudo-F = 8.289, p = 0.02, df = 29 respectively), however, PERMDISP analyses (p = 0.0069 and p = 0.002 respectively) revealed different initial dispersions amongst factor groups, which has an effect on the result of the PERMANOVA. Consequently, a dispersion effect is indeed present and perhaps a location effect as well (time and soil of origin factor influence) on the overall microbial composition. If the location of the samples in the PCO plot is inspected (Figure 2-4b) it is clear to the observer that soil of origin and time are most likely having a significant effect on the location of the samples and the overall microbial community.[3]

Table 2-7 shows top ten taxa for each allotment soil for which abundance increases in samples exposed to metaldehyde in the laboratory versus non-exposed controls at time 96 d. A heterogeneous group of enriched organisms was found for both soils in which the presence of the genera *Clostridium* and *Sphingomonas* in the lists from both soils stands out.

2.3.6 Growth of isolated strains using metaldehyde as a sole source of carbon

Growth of all the isolated degrading strains (and the previously isolated *A. calcoaceticus* E1) on metaldehyde (150 mg L^{-1}) as the sole source of carbon was achieved, and a strong correlation with the disappearance of metaldehyde was observed, supporting the conclusion that they use metaldehyde as a carbon and energy source (Figure 2-5). *A. bohemicus* JMET-C stood out as the strain with the shortest lag phase, doubling time, time required for metaldehyde removal and the highest maximum compound degradation rate (Table 2-8).

[3] A non-significant PERMDISP result is not strictly necessary to be able to interpret a PERMANOVA result, since PERMDISP will detect differences in dispersion that may not be significant enough to change the overall result of the PERMANOVA test (Anderson et al. 2008).

Table 2-7. Identification of top ten taxa for each allotment soil for which abundance increases in samples with a repeated exposure to metaldehyde in the laboratory versus non-exposed controls at time 96 d.

Soil	Family	Genus	Species	Fold increase
S-Met	Streptomycetaceae	*Streptomyces*	Unassigned	170
	Clostridiaceae	*Clostridium*	*subterminale*	161
	Microbacteriaceae	Unassigned	Unassigned	154
	Rhizobiaceae	*Rhizobium*	Unassigned	150
	Paenibacillaceae	*Thermobacillus*	Unassigned	137
	Pseudonocardiaceae	*Pseudonocardia*	*halophobica*	131
	Geodermatophilaceae	Unassigned	Unassigned	116
	Bacillaceae	Bacillus	flexus	114
	Micromonosporaceae	*Dactylosporangium*	Unassigned	94
	Sphingomonadaceae	*Sphingomonas*	Unassigned	85
S-NoMet	Xanthomonadaceae	*Lysobacter*	Unassigned	195
	Pseudonocardiaceae	*Saccharomonospora*	Unassigned	125
	Erythrobacteraceae	*Erythrobacteraceae*	Unassigned	106
	Burkholderiaceae	Unassigned	Unassigned	102
	Clostridiaceae	*Clostridium*	Unassigned	95
	Sphingomonadaceae	*Sphingomonas*	*wittichii*	69
	Koribacteraceae	Unassigned	Unassigned	59
	Rhodospirillaceae	*Stella*	vacuolata	48
	Sporichthyaceae	*Sporichthya*	Unassigned	43
	Gracilibacteraceae	Unassigned	Unassigned	42

Samples from S-Met and S-NoMet soils were both rarefied to 64225 reads.

The rest of the *Acinetobacter* strains followed with respect to the time needed for metaldehyde elimination; *P. vancouverensis* SMET-B came next. The *Sphingobium* CMET-H/HMET-G strains were intermediate in this regard, while the *Rhodococcus globerulus* HMET-A/HNO-A and *C. jiangsuensis* SNO-D strains were slowest.

Table 2-8. Bacterial growth and metaldehyde degradation statistics for the isolated strains growing on MSM with metaldehyde (150 mg L^{-1}) as the sole source of carbon

Strain	Code	Lag phase (h)	Doubling time (h)	Peak OD at 600 nm	Max. degradation rate (mg L^{-1} h^{-1})	Time for degr. below LD (h)
Acinetobacter calcoaceticus	E1	12	3.11	0.149	14.85	20
Acinetobacter bohemicus	JMET-C	8	1.48	0.323	31.96	15
Sphingobium sp.	CMET-H	21	4.59	0.210	11.31	42
Rhodococcus globerulus	HMET-A	29	2.84	0.263	13.74	44
Sphingobium sp.	HMET-G	19	3.07	0.187	14.87	35
Rhodococcus globerulus	HNO-A	29	3.28	0.274	15.71	44
Pseudomonas vancouverensis	SMET-B	15	3.04	0.213	17.02	26
Acinetobacter lwoffii	SMET-C	11	1.64	0.229	17.48	24
Caballeronia jiangsuensis	SNO-D	26	5.08	0.201	11.65	47

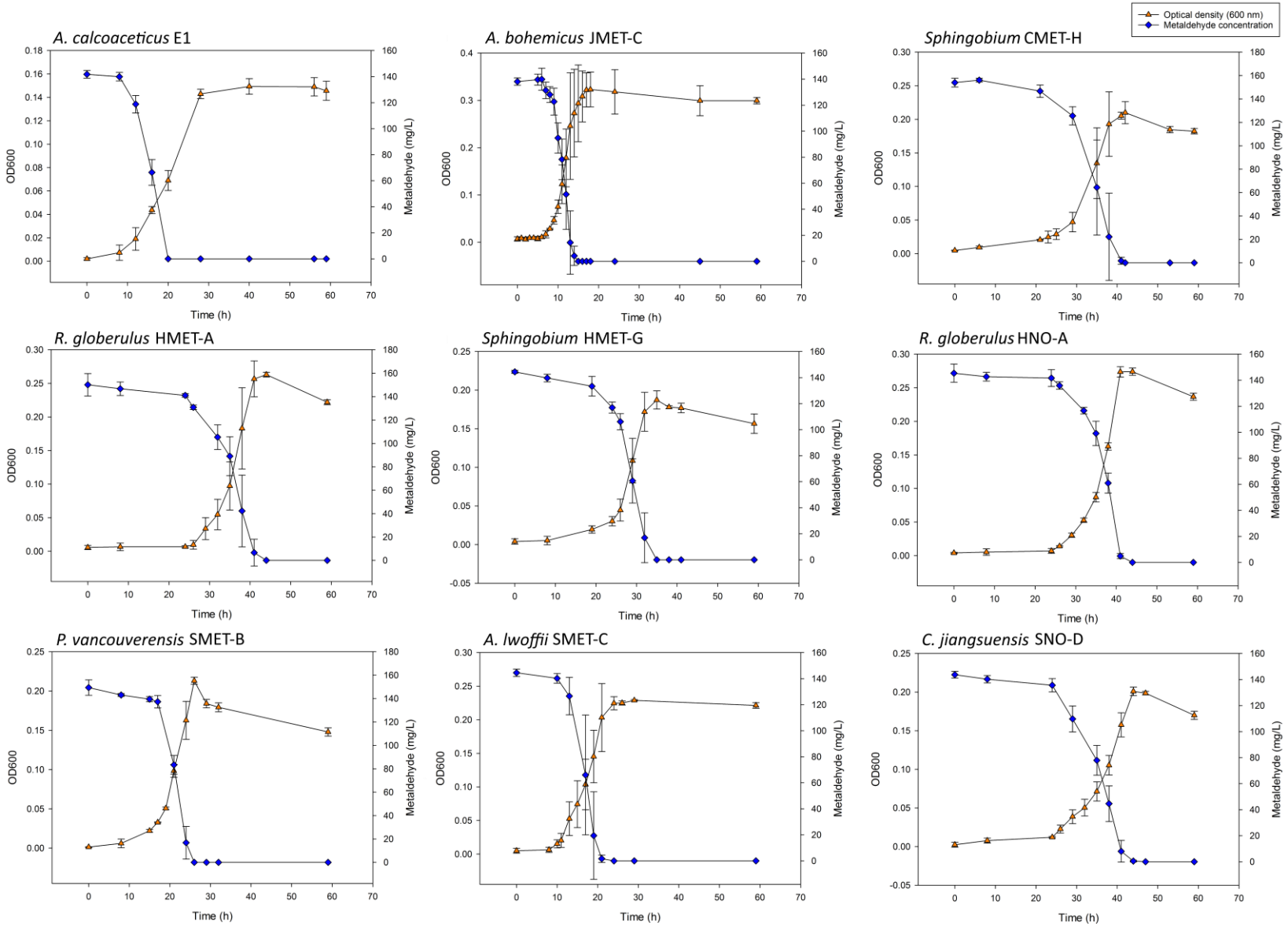


Figure 2-5. Growth of degrading strains on metaldehyde (150 mg L^{-1}) as the sole source of carbon. The inoculum was substituted by an equal volume MSM in the abiotic controls. Bars represent standard deviation for 3 replicates. Final metaldehyde concentration in the abiotic controls was 99.9% (CV = 3.0%) of the starting concentration.

2.3.7 Comparative genomic analysis of metaldehyde-degrading strains

Diverse metaldehyde-degrading strains were isolated, and a draft genome for each one was obtained. Comparative genomic analyses focused on identifying shared genes encoding proteins needed for metaldehyde degradation.

A previous study had identified an *Acinetobacter* strain (E1) capable of degrading metaldehyde (Thomas et al. 2017); additionally, two other strains of *Acinetobacter* (*A. bohemicus* JMET-C and *A. lwoffii* SMET-C) were isolated in this study. The predicted proteomes from these strains, along with those of the metaldehyde-degrading strains *P. vancouverensis* SMET-B, *C. jiangsuensis* SNO-D, and the non-degrading strain *A. calcoaceticus* RUH2202 were compared through Blast Score Ratio (BSR) analysis using *A. calcoaceticus* E1 as reference. A total of 65 candidate proteins were identified as present in all metaldehyde-degrading strains of *Acinetobacter* but absent from *A. calcoaceticus* RUH2202. Inclusion of the genomes from more evolutionarily divergent strains SNO-D and SMET-B decreased the number of predicted proteins shared amongst the metaldehyde-degrading strains and absent from RUH2202 to four (Table 2-9). Identical results were obtained when the analysis was repeated using *A. bohemicus* ANC3994 as non-degrading strain. A graphical representation of the results is displayed in Figure 2-6.

Table 2-9. Predicted proteins shared between metaldehyde-degrading strains, absent from non-degrading strains and their respective BSR values when compared against the reference genome (*A. calcoaceticus* E1).

	MahX		MahY		AldH		TnpA	
	aa length	BSR[1]	aa length	BSR[1]	aa length	BSR[1]	aa length	BSR[1]
A. calcoaceticus E1	314	1.000	149	1.000	231	1.000	503	1.000
A. calcoaceticus RUH2202	NP[2]	-	NP[2]	-	495	0.441	NP[2]	-
A. bohemicus JMET- C	314	1.000	149	1.000	231	1.000	503	1.000
A. lwoffii SMET-C	314	0.994	149	0.993	231	1.000	387	0.997
P. vancouverensis SMET-B	314	0.975	149	0.993	39	1.000	327	0.997
C. jiangsuensis SNO-D	314	0.984	149	0.993	88	1.000	262	1.000

[1]BSR: Blast-Score Ratio
[2]NP: not present

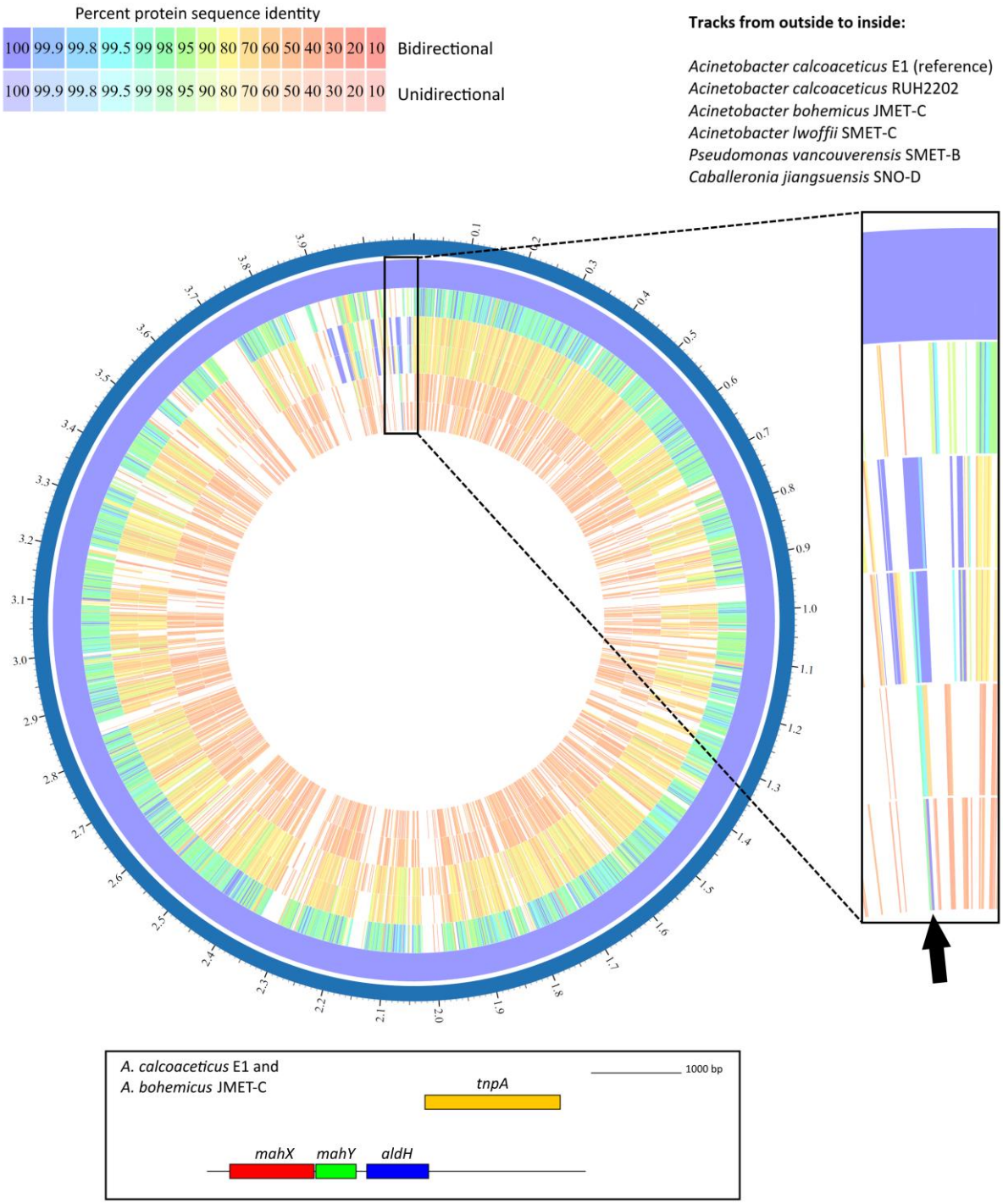


Figure 2-6. Proteome comparison between metaldehyde-degrading and non-degrading strains using *Acinetobacter calcoaceticus* E1 as reference. The analysis was performed using PATRIC Proteome Comparison Service (Wattam et al., 2017). A diagram of the shared metaldehyde-degrading gene cluster is shown in the inset. *mahX*: 2-oxoglutarate and Fe(II)-dependent oxygenase; *mahY*: vicinal oxygen chelate protein; *aldH*: NAD(P)$^+$-dependent aldehyde dehydrogenase; *tnpA*: Y2-transposase.

The identified proteins corresponded to a single cluster of four apparently horizontally-transferred genes, judging by the high similarity between the degrading proteins of the different isolates reflected in the high BSR results. This hypothesis will be explored in greater detail in Chapter 3. Predicted protein sequences were used to search NCBI's conserved domain database (Marchler-Bauer et al. 2015), using the cluster from *A. bohemicus* JMET-C as query. Thresholds of 50% and 70% identity have been proposed for assignment of third and full Enzyme Commission numbers at the domain level (Addou et al. 2009), and these values were used to support predicted functional assignment and naming of genes. Where only lower sequence identity to existing genes was observed, genes were denoted *mah* (for *m*et*a*lde*h*yde). The first gene in the cluster (*mahX*) contains a main domain classified into the 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily, with the highest sequence identity (49%) to a protein involved in biosynthesis of mitomycin antibiotics from *Sphingobium japonicum*. The second gene (*mahY*) is related to the vicinal oxygen chelate family, which is found in a variety of structurally related metalloproteins, including the type I extradiol dioxygenases, glyoxalase I and a group of antibiotic resistance proteins, with the highest sequence identity (41%) to a hypothetical unannotated protein from *Mycolicibacterium moriokaense*. The third gene (*aldH*) could be confidently assigned as an $NAD(P)^+$-dependent aldehyde dehydrogenase due to high identity (78%) to an aldehyde dehydrogenase from *Solimonas* sp. The fourth gene (*tnpA*), annotated as Y2-transposase, is almost identical (99%) to a transposase from *Pseudomonas pseudoalcaligenes*. All these functions are consistent with a potential horizontally-transferable metaldehyde degradation pathway. In *P. vancouverensis* SMET-B the *tnpA* was found to be located elsewhere in the genome. In *P. vancouverensis* SMET-B and *C. jiangsuensis* SNO-D, *aldH* was truncated.

2.4 Discussion

Previous work has indicated microbial activity is involved in degradation of the xenobiotic pesticide metaldehyde (Simms et al. 2006; Thomas et al. 2017). Here a systematic molecular and microbiological approach has been used to analyse and improve isolation techniques, and gain insight into the abundance, distribution, and mechanisms of microbial metaldehyde degradation.

The isolation of diverse metaldehyde-degrading strains was successfully combined with increasingly affordable whole-genome sequencing to, by means of comparative genomics, identify a xenobiotic-degrading gene cluster. This strategy has proven effective for other xenobiotics in the past as well (Yan et al. 2016), and as whole-genome sequencing becomes increasingly affordable, it has the potential of becoming a very important approach to identifying organic compound-degrading gene clusters. This functional assignment was subsequently confirmed by E Fuller using chemical mutagenesis and heterologous expression of the enzyme (Castro-Gutiérrez et al. 2020). This illustrates the value of obtaining a broad collection of isolates for the identification of degrading mechanisms.

Availability of labile carbon is considered the main limiting factor for microbial growth in soil (Aldén et al. 2001), and metaldehyde, when added to the soil, constitutes such a source of carbon. In this context, the presence of genes for its degradation provide the host with the ability to utilize this readily available carbon source, providing a selective advantage. Genes *mahX* and *mahY* have a moderate similarity to other well-characterized genes, so they may share an evolutionary ancestor with them. However, they appear to have diverged sufficiently so that new catalytic properties, such as substrate specificity towards metaldehyde, has emerged, was selected for, and transferred to other hosts via transposable elements in plasmids or other vectors. The Moir Group (J. Moir, E. Fuller, V. Castro-Gutierrez) has proposed a catabolic pathway for metaldehyde based on their probable activity. *mahX* encodes a predicted 2-oxoglutarate-dependent oxygenase. Given that this gene is sufficient to bring about at least the

initial step of metaldehyde degradation, it is proposed that protein MahX is an oxygenase that activates metaldehyde metabolism by oxygenation and ring cleavage. The predicted product is a hemiacetal (1,3,5,7-tetramethyl-2,4,6-trioxa-1-hydroxy-7-octanone) (Figure 2-7). Whilst this hemiacetal is unstable, the timescale of its chemical degradation is likely to be minutes to hours (Chiang and Kresge 1985). It can be speculated that MahY acts as a lyase that accelerates the iterative breakdown of the hemiacetal intermediate to acetaldehyde (Figure2-7). MahY is most closely related to the vicinal oxygen chelate (VOC) superfamily (He and Moran 2011). Whilst the predicted substrate here is not a VOC *per se*, VOC family members bind substrates with two oxygen atoms and some members of the family are lyases in keeping with the predicted function here. Acetaldehyde generated from MahY is predicted to be converted to acetate by AldH (Fig. 2-7), and subsequently incorporated into central metabolism.

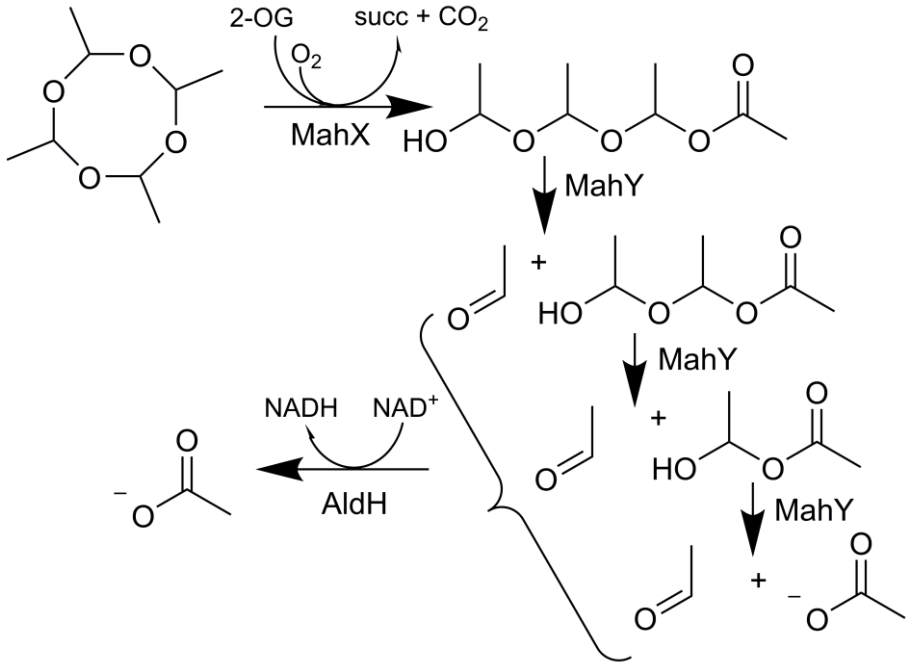


Figure 2-7. Predicted pathway for metaldehyde degradation. MahX is related to 2-oxoglutarate (2-OG)-dependent oxygenases that generate succinate (succ) and CO_2. MahX oxygenates metaldehyde to release a linear hemiacetal that is cleaved iteratively into acetaldehyde + a shorter chain hemiacetal, and eventually acetate. AldH oxidises acetaldehyde to acetate in an NAD^+-dependent reaction. Figure by J. Moir.

Since the degrading gene cluster identified in *Acinetobacter*, *Caballeronia* and *Pseudomonas* is not present in some of the other metaldehyde-degrading isolates from this study (*Sphingobium* sp. CMET-H and HMET-G, *R. globerulus* HMET-A and HNO-A) and from an earlier study (*Variovorax* sp. E3) (Thomas et al. 2017), as shown by whole-genome sequencing and PCR-based gene detection (data not shown), it is clear that additional metaldehyde-degrading mechanisms are found in nature.

Previous attempts to isolate metaldehyde-degrading strains from soils resulted in varying degrees of success. Thomas and collaborators (2017) isolated two metaldehyde-degrading strains from domestic soils, nevertheless, additional attempts using three different previously-exposed agricultural soils failed. Initially, similar results were obtained; isolation of degraders was occasional and only from previously exposed soils. Even though previous applications of metaldehyde to the soil in the field seemed to aid in the subsequent isolation of degraders, it was far from a guarantee. Many variables such as the length of time between application and soil sampling, the dose of pesticide, its formulation (and thus distribution and fate in the field), and the sampling regime, may all influence the abundance of metaldehyde degraders in the soil samples, and thus the success of isolation strategies. Furthermore, laboratory culture conditions cannot fully replicate the ideal conditions to suit the physiology of many of members of the soil microbial community.

The fact that metaldehyde is normally applied as a pellet in the field (vs. liquid spray forms) influences its behaviour and distribution in soil (Bond 2018). This is important because a gradient of decreasing concentrations in soil would be expected as distance from the pellet increases. The input of labile C sources to soil increases the abundance and activity of microorganisms generating microbial hotspots (Kuzyakov and Blagodatskaya 2015). The distribution of metaldehyde after pellet applications may generate defined zones of enhanced metaldehyde degradation in soil. This would pose a challenge when sampling, because high biodegradation hotspots, with elevated numbers of degraders, could easily be missed.

Consistent isolation of metaldehyde degraders was possible from horticultural soils, regardless of their metaldehyde exposure history, after adding metaldehyde to soil samples in the laboratory and incubating them for a defined period. This approach minimized the challenges to isolation that would arise from distribution of activity in hotspots. Further analysis revealed that metaldehyde addition increased the number of culturable metaldehyde degraders in the samples, while incubation in laboratory conditions (initial homogenization, 25°C, 100% relative humidity) also increased the total bacterial biomass, greatly facilitating, in combination, the isolation of degraders. Even though this approach has been used in the past for other pesticides (Goda et al. 2010; Perruchon et al. 2016; B. Singh et al. 2004) it seems to be of particular importance for the isolation of metaldehyde degraders because the uneven distribution of the pesticide generated from normal field applications makes the direct isolation from freshly collected soils difficult.

In this study metaldehyde degradation occurred faster in soils previously exposed to metaldehyde in the field compared to soils with very similar physicochemical characteristics but not directly exposed to the pesticide. Furthermore, even a single metaldehyde addition to soil samples in the laboratory led to accelerated degradation of metaldehyde and accumulation of culturable metaldehyde-degrading microbes. This illustrates that metaldehyde degrading strains are strongly selected by metaldehyde in soil. In line with this observation, Balashova et al. (2020) saw an significant increase in catabolic competence in many soils following slug pellet addition.

Culture-independent analysis during metaldehyde degradation showed that the main factors governing overall bacterial community composition were soil origin and incubation time, not metaldehyde. Several studies have found important pesticide-driven changes in community structure with other pesticides only when using doses of several times the recommended application dose or after several repeated additions (Crouzet et al. 2010; Cycoń et al. 2013; Itoh et al. 2014). Nevertheless, at a finer level, several specific taxa were enriched after metaldehyde application (Tables 2-6 and 2-7). Notably, the lists of most enriched taxa in all soils after

metaldehyde application from the first round of assays were dominated by anaerobes, which may not reflect the specific groups executing metaldehyde degradation in soil but may instead be related to increased bacterial metabolism and abundance, leading to anaerobic conditions developing in the soils during the incubation due to increased oxygen respiration (Shennan et al. 2018; Streminska et al. 2014). This is consistent with the fact that enrichment of specific anaerobic taxa was observed even in the already metaldehyde-adapted soil S-Met.

A metaldehyde-degrading *Sphingobium* was isolated following enrichment (in soil H-Met), and this genus was identified, through amplicon sequencing, as being enriched > 300-fold by metaldehyde in the soil (Table 2-6). Beyond that, no other cultured metaldehyde-degraders were identified as enriched in soil amplicon sequencing. This is presumably a result of metaldehyde degradation being a rare trait, possessed by a relatively low proportion of the total microbial community (reflected by 10^3-10^5 metaldehyde-degraders (Fig. 3a) out of an estimated 10^8-10^9 microbial cells/g of soil in our samples).

2.5 Conclusions

This chapter highlights the continuing value of "traditional" experimental enrichment methods for obtaining a large library of degrading strains and their application in coordination with contemporary molecular methods to elucidate the gene cluster responsible for metaldehyde degradation. An insight into the diversity and mechanism of biological metaldehyde degradation has been provided in this chapter; this understanding will be essential for predicting the environmental fate of the compound and for optimizing and monitoring the performance of engineered biological systems for metaldehyde removal from drinking water.

CHAPTER 3: PRESENCE, PROLIFERATION AND DISTRIBUTION OF BACTERIAL METALDEHYDE-DEGRADING GENES IN SOILS FROM NORTHERN ENGLAND

3.1 Introduction

The assessment and prediction of the fate of xenobiotics in the environment during bioremediation or natural attenuation can be carried out through different strategies: analysing the disappearance of the parent compound, detecting its transformation products or through evidence of transformation potential in a given setting (Fenner et al. 2013). It is in this third approach in which knowledge regarding degrading genes can be used to detect and quantify their copy numbers or their expression, which provides a measure of degradation activity.

Several approaches for detection and quantification of specific xenobiotic degrading strains in environmental samples through marker (non-degrading) genes have been undertaken in the past (Widada et al. 2002). However, since pesticide degrading genes can be acquired by previously non-degrading microbes (DiGiovanni et al. 1996) or lost from degrading organisms (Changey et al. 2011), detecting and quantifying the degrading genes themselves carries the obvious advantage of more accurately reflecting the status of the degrading community.

In the previous chapter a culture-based MPN technique was used to quantify metaldehyde degraders in soil samples. Afterwards, the identification of a single shared cluster of three functional genes (*mahX*, *mahY* and *aldH*) responsible for the first steps of metaldehyde degradation in bacteria from different taxonomic origins was achieved. The recently acquired knowledge regarding these degrading genes makes it now possible to design molecular strategies for their detection in environmental samples to assess pesticide-degrading potential and activity, an objective that is undertaken in this chapter.

Moreover, horizontal gene transfer of this cluster was hypothesized due to the high similarity of the protein sequences between isolates (Chapter 2). However further evidence of

horizontal gene transfer had not been explored and a specific genetic element responsible for carrying and dispersing this trait had not been identified and analysed.

The increasing affordability of newer and more informative sequencing approaches has allowed environmental microbiologists to gain better insight into the wider genetic and ecological context in which these degrading genes and organisms are immersed. Third generation long-range sequencing approaches have greatly improved *de novo* genome assembly and analysis of genome structure (Lee et al. 2016). By using these techniques, the identification and description of specific replicons and insertion sequences responsible for maintaining and dispersing metaldehyde-degrading genes in the soil environment was carried out as described in this chapter.

3.2 Materials and methods

3.2.1 Chemicals and reagents

Metaldehyde (99%) was purchased from Acros Organics, NJ; sodium sulfate (≥99%) was purchased from Honeywell-Fluka, Bucharest; all other chemicals were purchased from Sigma-Aldrich, St. Louis, MO.

3.2.2 Combined Illumina / Oxford Nanopore sequencing for metaldehyde degrader *A. calcoaceticus* E1

Combined Illumina and Oxford Nanopore sequencing was undertaken with the aim to obtain the complete and fully assembled genome sequence of *Acinetobacter calcoaceticus* E1 (Thomas et al. 2017), which has been used so far as our reference strain for metaldehyde degradation. The sequencing service was provided by MicrobesNG (Birmingham, UK). The strain was grown to late exponential phase in LB broth (30°C, 150 rpm), cells were gently pelleted down (10 min, 500 g), supernatant removed, cells resuspended in cryopreservative fluid with a wide

bore tip and stored in a provided Microbank bead tube (Pro-Lab Diagnostics UK, United Kingdom).

In the MicrobesNG protocol for Illumina short-read sequencing, three beads were washed with extraction buffer containing lysozyme and RNase A, incubated for 25 min at 37°C. Proteinase K and RNaseA were added and incubated for 5 min at 65°C. Genomic DNA was purified using an equal volume of SPRI beads and resuspended in EB buffer. DNA was quantified in triplicates with the Quantit dsDNA HS assay in an Ependorff AF2200 plate reader. Genomic DNA libraries were prepared using Nextera XT Library Prep Kit (Illumina, San Diego, USA) following the manufacturer's protocol with the following modifications: two nanograms of DNA instead of one were used as input, and PCR elongation time was increased to 1 min from 30 seconds. DNA quantification and library preparation were carried out on a Hamilton Microlab STAR automated liquid handling system. Pooled libraries were quantified using the Kapa Biosystems Library Quantification Kit for Illumina on a Roche light cycler 96 qPCR machine. Libraries were sequenced on the Illumina HiSeq using a 250bp paired end protocol. Reads were adapter trimmed using Trimmomatic 0.30 with a sliding window quality cut-off of Q15 (Bolger et al. 2014).

For Oxford Nanopore sequencing, approximately $2x10^9$ cells were used for high molecular weight DNA extraction using Nanobind CCB Big DNA Kit (Circulomics, Maryland, USA). DNA was quantified with the Qubit dsDNA HS assay in a Qubit 3.0 (Invitrogen) Eppendorf UK Ltd, United Kingdom. Long read genomic DNA libraries were prepared with Oxford Nanopore SQK-RBK004 kit and SQK-LSK109 kit with Native Barcoding EXP-NBD104/114 (ONT, United Kingdom) using 400-500 ng of HMW DNA. Twelve to twenty-four barcoded samples were pooled together into a single sequencing library and loaded in a FLO-MIN106 (R.9.4) flow cell in a GridION (ONT, United Kingdom).

3.2.3 Bioinformatic analyses of whole-genome sequencing data

De novo genome assembly was performed using Unicycler v0.4.0 (Wick et al. 2017) and contigs were annotated using Prokka 1.11 (Seemann 2014). Replication origin location was determined using the DoriC replication origin database (Luo and Gao 2019). GC content and GC skew were determined using Artemis 18.1.0 (Carver et al. 2012). rRNA operons were identified using RNAmmer 1.2 (Lagesen et al. 2007). tRNA genes were identified using tRNAscan-SE (Lowe and Chan 2016). Additional functional annotations were performed using eggNOG-mapper against the EggNOG V 5.0 database (Huerta-Cepas et al. 2017, 2019). Additional promoter predictions were performed using BPROM (Solovyev et al. 2011). Plasmid mobilization proteins and associated regions were determined using the OriT finder server (X. Li et al. 2018). Searches against the ISFinder database (Siguier et al. 2006) were used to characterize transposases and insertion sequences. Protein comparison between multiple sequences was performed and associated plots were generated using the PATRIC Proteome Comparison Service (Wattam et al. 2017). The Burrows-Wheeler Alignment tool (BWA) (Li and Durbin, 2009) in UGENE (Golosova et al., 2014; Okonechnikov et al., 2012) was used to map short DNA sequencing reads against specific scaffolds. PlasmidSPAdes 3.14.0 (Antipov et al., 2016) was used to determine the putative gene composition and structure of plasmids from short read DNA sequencing. Single sequence nucleotide and protein comparisons were carried out using nBLAST or pBLAST searches (Altschul et al. 1990). Plots were generated using Snapgene and DNAplotter (Carver et al. 2009).

3.2.4 Development of a qPCR assay for quantification of metaldehyde-degrading genes

3.2.4.1 Sequence alignment and primer design

Three functional genes (*mahX*, *mahY*, *aldH*) and a transposase were found to be present in the gene cluster responsible for metaldehyde degradation. Sequences for the genes *mahX* and *mahY* were compared between *A. calcoaceticus* E1, *A. bohemicus* JMET-C, *A. lwoffii* SMET-

C, *P. vancouverensis* SMET-B and *C. jiangsuensis* SNO-D. Coding regions for each one of these genes were aligned using MUSCLE (Edgar, 2004) in MEGA7 (Kumar et al., 2015). Conserved regions were identified and selected as templates for primer design. Base substitution quantification was conducted using the Maximum Composite Likelihood model (Tamura et al. 2004). Primers for qPCR analysis were designed using the PrimerQuest online tool (Integrated DNA Technologies) with optimal parameters set to Tm= 62°C, GC%= 50, primer size= 22 bp, amplicon size= 100 bp. Absence of predicted secondary structure formation and primer dimers was corroborated *in silico.* Specificity of the primers was first checked *in silico* through BLAST searches against the entire NCBI nucleotide collection. No target sequences with significant similarity to both the forward and reverse primers capable of generating an amplicon were detected. The list of the designed qPCR primers is shown in Table 3-1. All primers were synthesized by Sigma.

Table 3-1. qPCR primer sequences for genes in the metaldehyde-degrading cluster

Target gene	Oligonucleotide	Sequence (5'-3')	Location*	Amplicon length (bp)	Tm (°C)
mahX	mahX-86F	GTCAGCCGGACACTGATTT	86-104	98	62
mahX	mahX-183R	CGGAGGAACCACGCAATAG	165-183		62
mahY	mahY-143F	GGCTCGGTGTCGAACTTATT	143-162	145	62
mahY	mahY-287R	AGGTGCTCGATGTCAGATTTC	267-287		62

*Nucleotide position in reference to *A. calcoaceticus* E1

3.2.4.2 Initial testing through endpoint PCR

Primers were initially tested using endpoint PCR visualizing the products on 1.2% agarose gels with different number of cycles (30-40) and a temperature gradient (50-65°C) to establish the optimal annealing temperature for each primer pair. Extracted DNA from a collection of metaldehyde-degrading strains with and without the degrading gene cluster (as determined by whole-genome sequencing) and pre-exposed soils from the first round of metaldehyde degradation studies (soils S-Met and H-Met) were used for this purpose.

The optimal thermocycling conditions for both primer pairs mahX-86F/183R and mahY-143F/287R in endpoint PCR were: 95 °C/3 min; 95 °C/30 s, 60 °C/30 s and 72 °C/30 s (35 cycles), and 72 °C/5 min. PCR reaction mixture (50 µL) contained 2 mM $MgCl_2$, 200 µM each dNTP, 1.25 U Taq-polymerase, 400 nM each primer, 20 µg BSA and 1 ng of template DNA (from strains) or 40 ng of template DNA (from soils).

3.2.4.3 Generation of quantification standards for qPCR

To generate quantification standards for qPCR, primers amplifying a broader region of genes *mahX* and *mahY* were designed in a similar way as described above but setting optimal amplicon size to 250 bp. Primer sequences for quantification standards are shown in Table 3-2.

Table 3-2. PCR primer sequences for generating quantification standards

Target gene	Oligonucleotide	Sequence (5'-3')	Location*	Amplicon length (bp)	
mahX	mahX-10F	GAGCTTGAGTCCGCCGTGAA	10-29	244	60
mahX	mahX-253R	TGCGTCTTTCCGCGATCTCC	234-253		60
mahY	mahY-69F	GATCTTCGCCGTCCCTTTC	69-87	260	62
mahY	mahY-328R	AAGAGAAGGTCGGTCCGTAT	309-328		62

*Nucleotide position in reference to *A. calcoaceticus* E1

Similarly, these primers were tested using endpoint PCR visualizing the products on 1.2% agarose gels with a temperature gradient (50-65°C) to establish the optimal annealing temperature for each primer pair. The optimal thermocycling conditions for primer pair mahX-10F/253R were: 95 °C/3 min; 95 °C/30 s, 64 °C/30 s and 72 °C/30 s (30 cycles), and 72 °C/5 min. For mahY-69F/328R these were: 95 °C/3 min; 95 °C/30 s, 58 °C/30 s and 72 °C/30 s (30 cycles), and 72 °C/5 min. PCR reaction mixture (50 µL) contained 2 mM $MgCl_2$, 200 µM each dNTP, 1.25 U Taq-polymerase, 400 nM each primer, 20 µg BSA and 1 ng of template DNA, for which genomic DNA from *A. bohemicus* JMET-C was used.

PCR products were purified using QIAGEN PCR purification kit and quantified using Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific) according to manufacturer's protocol to generate

the quantification standards for qPCR. Standard curves using each of the purified PCR standards were generated from 10^1 to 10^9 target copies per reaction. All dilutions were performed in loBind DNA tubes (Eppendorf, Hamburg, Germany).

3.2.4.4 Intercalating dye-based qPCR

Intercalating dye qPCR was performed for genes *mahX* and *mahY* in 20 µL reactions containing 10 µL FastSybr Green Master Mix (2x), 0.4 µL BSA (20 mg/mL), 1 µL each primer (10 µM), 7.8 µL nuclease free water, 1 µL template DNA. All standard levels and samples were tested in triplicate. qPCR was performed in a QuantStudio 3 thermal cycler (Thermo Fisher Scientific). Thermocycling conditions were 95 ºC/20 s; 95 ºC/1 s, 60 ºC/20 s (40 cycles). A melting curve was added at the end to check for unspecific amplification: 95 ºC/1 s, 60 ºC/20 s, and from there increasing to 95 ºC again at a rate of 0.1 ºC/s. After the qPCR, amplification products were run in 1.2% agarose gels to verify the melting curve results. Subsequent analysis of the results showed that even though the assay had a high efficiency, sensitivity was lower than what was needed for soil samples; therefore, this issue was addressed by using probe-based qPCR instead of intercalating dye qPCR.

3.2.4.5 Probe-based qPCR

To increase the specificity and sensitivity of the assay, a qPCR method using a nucleotide probe instead of intercalating dyes for signal production was developed for gene *mahY*. The nucleotide probe was designed using the PrimerQuest tool with optimal parameters set to Tm= 68°C, GC%= 50, primer size= 24 bp. Probe information is presented in Table 3-3. The oligonucleotide probe was synthesized by Sigma.

Table 3-3. Nucleotide probe for gene *mah*Y

Target gene	Oligonucleotide	Sequence (5'-3')	Location*
*mah*Y	probe177	[6FAM]CAGCGACAGCCAGTTTGCTCAGGA[BHQ1]	177-200

*Nucleotide position in reference to *A. calcoaceticus* E1

For probe-based qPCR, each 20 µL reaction contained 10 µL inhibitor-resistant KAPA Probe Force Master Mix (2x), 1 uL mahY-probe177 (10 uM), 0.5 µL BSA (20 mg/mL), 1.8 µL each primer (10 µM), 3.9 µL nuclease free water, 1 µL template DNA. All standard levels and samples were tested in triplicate. qPCR was performed in a QuantStudio 3 thermal cycler (Thermo Fisher Scientific). Thermocycling conditions were 98 ºC/3 min; 95 ºC/10 s, 60 ºC/20 s (45 cycles).

Cycle thresholds (Ct) where sample fluorescence exceeds background fluorescence were recorded for the samples and quantification standards. The numbers of *mahX* or *mahY* gene targets were interpolated from the standard curve generated from the quantification standards in relation to their Ct. Soil types for which none of the samples appeared to be above the detection limit of the technique were tested for inhibitors by combining DNA from a positive sample of a different soil with DNA from the negative samples at ratios of 2:1, 1:1, and 0.1:1 and checking if an effect on amplification was evident.

3.2.5 Soil sampling and characterization

Four different additional soils were collected in Northern England from allotment and agricultural plots for metaldehyde removal, microbial community analyses and degrading gene quantification. Details for each plot are shown in Table 3-4. For each plot, triplicate 300 g sub-samples were taken from the top 10 cm of the soil and stored in plastic bags. These were combined into a composite sample and stored in loosely tied plastic bags in a temperature-controlled room at 23°C until analysed one week later. Soil samples were air dried overnight and sieved (2 mm mesh). Physical and chemical parameters for soils were determined at YARA, UK using accredited methods.

Table 3-4. Details for the soil collections performed for metaldehyde removal, degrading gene and microbial community analysis.

Soil code	Collection date	Location	Coordinates	Type of plot	Crop	Metaldehyde use history
STA	June 2019	Urlay Nook Road Allotments, Eaglescliffe, Stockton on Tees, County Durham	54.520417, -1.369611	Allotment	Broad beans	Applied for three months before sampling only
PA	September 2019	Nafferton Farm, Stocksfield, Northumberland	54.981343, -1.901597	Agricultural	Wheat	Metaldehyde applied one year before sampling
ING	September 2019	Nafferton Farm, Stocksfield, Northumberland	54.986730, -1.911462	Agricultural	Oilseed rape	Metaldehyde applied one year before sampling
LE	September 2019	Leeds Farm, Tadcaster, Leeds, West Yorkshire	53.868944, -1.326139	Agricultural	Grass-clover mix	No metaldehyde application history

3.2.6 Analytical methods for metaldehyde quantification

Analytical methods for metaldehyde extraction and quantification in soils were described in detail in Chapter 2.

3.2.7 Metaldehyde degradation profiles and microbial community changes in soil microcosms

Metaldehyde degradation was followed in the laboratory in STA, PA, ING and LE soils after a single metaldehyde application at 15 mg kg^{-1}. Samples were taken for soil genomic DNA extraction to determine changes in the soil microbial community in response to metaldehyde addition and incubation. Controls with no metaldehyde addition were labelled cSTA, cPA, cING and cLE. The process was carried out in the same way as described in Chapter 2. The protocol for 16S rRNA gene amplicon sequencing data analyses was also described in detail in Chapter 2.

3.2.8 Quantification of the metaldehyde-degrading gene *mahY* in soils

DNA from soils S-Met, S-NoMet, H-Met and H-NoMet, (used in the previous chapter) and from soils STA, PA, ING, and LE, and their respective controls, were used for *mahY* gene quantification using the newly developed probe-based qPCR protocol.

3.2.9 Isolation of metaldehyde-degrading strains

Metaldehyde-degrading strains were isolated from soil STA only using the enrichment culture procedure described in detail in Chapter 2.

3.2.10 Accession numbers

Raw reads for 16S rRNA amplicons were deposited in the European Nucleotide Archive under study PRJEB40574. Whole-genome enhanced sequencing data for the metaldehyde-degrading strain *A. calcoaceticus* E1 was deposited under accession numbers SAMEA7370790 (Oxford Nanopore reads), SAMEA7370296 and SAMEA7370297 (Illumina MiSeq reads).

3.3 Results

3.3.1 Combined Illumina and Oxford Nanopore sequencing for *A. calcoaceticus* E1

Illumina short-read sequencing and Oxford Nanopore sequencing were used in combination to provide a better understanding of the genomic organization of metaldehyde degrader *A. calcoaceticus* E1. Quality statistics for Illumina short-read sequencing and preliminary assembly are shown in Table 3-5.

Table 3-5. Quality statistics for Illumina short read whole-genome sequencing run and assembly of metaldehyde-degrading *A. calcoaceticus* E1.

Median insert size (bp)	Mean coverage (fold)	Number of reads	Number of contigs	Largest contig (bp)	Total length (bp)	GC (%)	N50 (bp)	L50 (bp)	# N's
321	70.3	774 073	112	377 626	4 402 996	38.7	105 976	13	0

Oxford Nanopore sequencing (mean coverage was 51.8x) provided the long reads needed to be used as a scaffold for the Illumina sequencing data. The complete and fully assembled genome sequence was successfully obtained. The genome consists of a circular chromosome and two circular plasmids. The result of this combined approach is presented in Table 3-6.

Table 3-6. Resulting contigs from combined Illumina and Oxford Nanopore sequencing and assembly of metaldehyde-degrading strain *A. calcoaceticus* E1.

Contig	Length (bp)	Type of replicon*	Topology	GC content (%)
1	4 406 004	Chromosome	Circular	38.8
2	76 942	Plasmid	Circular	41.9
3	10549	Plasmid	Circular	40.5
Total	4 493 495	-	-	38.8

*Presence of *rep* genes was used to identify replicons different from the chromosome

3.3.1.1 Chromosome

The map of the 4.4 Mb chromosome is displayed in Figure 3-1. Automated annotation identified 4292 coding sequences (CDSs) in the main chromosome. The starting point for the sequence was chosen to be the first base pair of the chromosomal replication initiator protein (*DnaA*). A 682 bp origin of replication (*oriC*) region was located immediately preceding the starting point and showed a 98.2% sequence similarity to the *oriC* locus region of *A. calcoaceticus* strain CA16 (ORI97013474) according to the DoriC replication origin database (Luo and Gao 2019). The location of the *oriC* region clearly matches the sign change in the GC skew calculation (Figure 3-1), which also provides an approximate estimation of the position of the replication terminus at the opposite end of the chromosome. A prominent peak in GC content (50.3%) can be observed in the region located approximately between 830-855 kb corresponding to a protein with an 81.76% nucleotide similarity (90% coverage) to a large Immunoglobulin-like domain-containing protein from *Acinetobacter pittii* (WP_101664692).

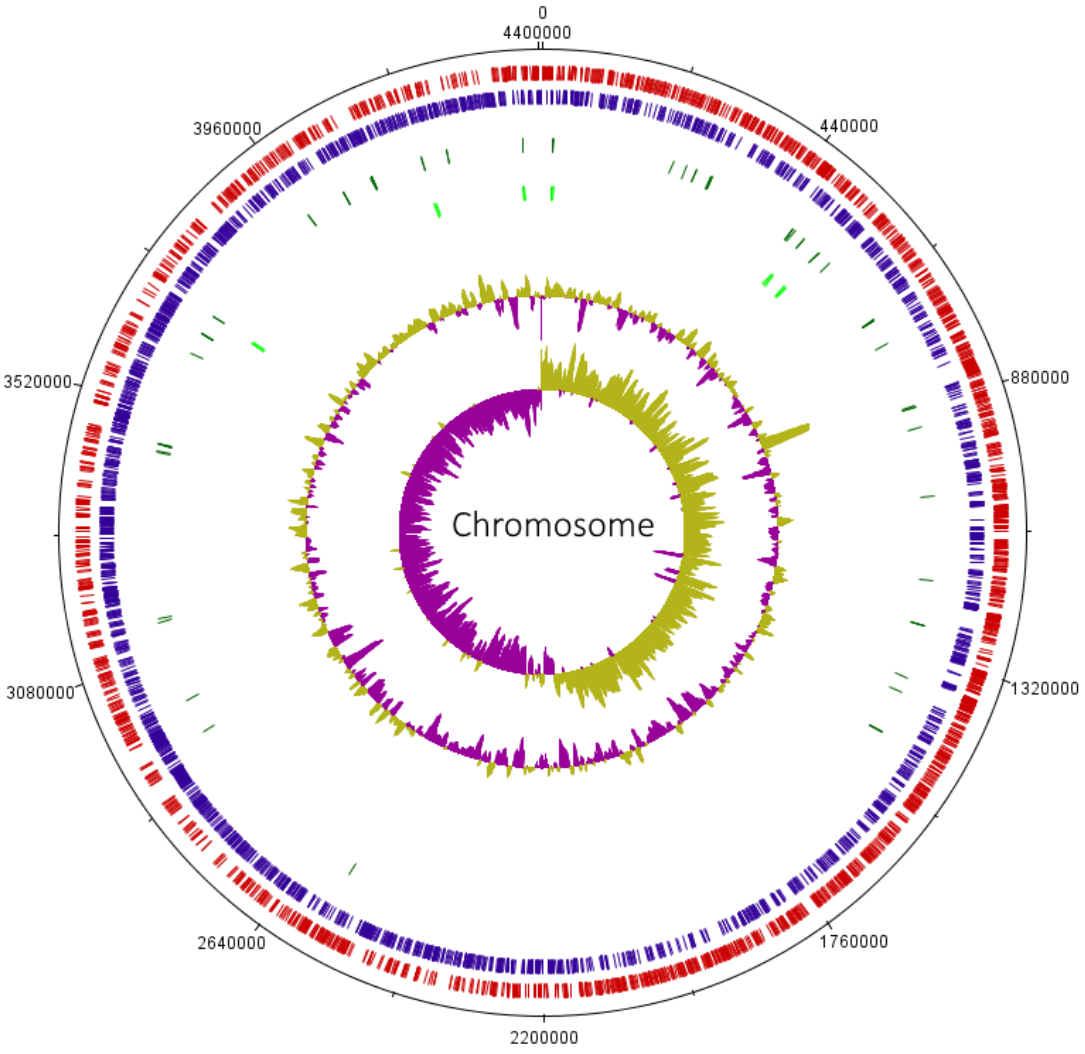


Figure 3-1. *A. calcoaceticus* E1 chromosome map. Sequence plot was generated using Artemis and DNAPlotter (Carver et al. 2009, 2012). From outside to inside: tick marks indicate 220 000 bp steps; open reading frames for the forward and reverse strands are indicated in red and blue; tRNA genes are indicated in dark green; rRNA genes are indicated in light green; GC content (window size = 10 000 bp, step size = 200 bp; purple: below average, gold: above average); innermost ring indicates GC skew (window size = 10 000 bp, step size = 200 bp; purple: below average, gold: above average).

Six rRNA operons were identified using RNAmmer (Lagesen et al. 2007), which is consistent with the mean for *Acinetobacter baumanni/calcoaceticus* genomes in rrnDB (Stoddard et al. 2015). 75 tRNA genes were also identified using tRNAscan-SE (Lowe and Chan 2016).

Functional annotation was performed using eggNOG-mapper against the EggNOG V 5.0 database (Huerta-Cepas et al. 2017, 2019) and the summary of functional categories is shown in Table 3-7.

Table 3-7. EggNOG functional categories for the predicted genes of the *A. calcoaceticus* E1 chromosome

Function	Number of genes
CELLULAR PROCESSES AND SIGNALING	
[D] Cell cycle control, cell division, chromosome partitioning	46
[M] Cell wall/membrane/envelope biogenesis	216
[N] Cell motility	37
[O] Post-translational modification, protein turnover, and chaperones	100
[T] Signal transduction mechanisms	96
[U] Intracellular trafficking, secretion, and vesicular transport	118
[V] Defense mechanisms	43
[W] Extracellular structures	1
[Y] Nuclear structure	0
[Z] Cytoskeleton	0
INFORMATION STORAGE AND PROCESSING	
[A] RNA processing and modification	2
[B] Chromatin structure and dynamics	0
[J] Translation, ribosomal structure and biogenesis	188
[K] Transcription	341
[L] Replication, recombination and repair	223
METABOLISM	
[C] Energy production and conversion	212
[E] Amino acid transport and metabolism	332
[F] Nucleotide transport and metabolism	95
[G] Carbohydrate transport and metabolism	167
[H] Coenzyme transport and metabolism	135
[I] Lipid transport and metabolism	197
[P] Inorganic ion transport and metabolism	256
[Q] Secondary metabolites biosynthesis, transport, and catabolism	122
POORLY CHARACTERIZED	
[R] General function prediction only	0
[S] Function unknown	835

3.3.1.2 Plasmid pAME76

Automated annotation identified 81 CDSs in the 76.9 kb circular plasmid, from here on called pAME76, of which 24 are transposases, 12 are hypothetical proteins with no annotation,

10 are involved in benzoate compound catabolism, 6 are part of toxin-antitoxin plasmid stability systems, 6 are related to mercuric resistance, 5 are integrases or recombinases, 3 are the metaldehyde-degrading proteins (including *aldH*), 2 are cation transporters, 1 is a replication initiation protein, 1 is a DNA polymerase and 11 encode for various other functions. Plasmid map is shown in Figure 3-2. No oriT region, relaxase gene, type IV coupling proteins and type IV secretion systems were found using OriT finder (X. Li et al. 2018), which implies that either pAME76 is a non-mobilizable plasmid, or could be due to a limitation in the database.

The three previously identified genes from the metaldehyde-degrading cluster (*mahX*, *mahY* and *aldH*) and an IS91 family transposase are flanked by two IS6 family transposases. Upon closer inspection using ISFinder (Siguier et al. 2006) it was determined that these IS6 transposases are two identical copies of an intact IS*Our*1 insertion sequence first described in *Oligella uretralis* (AY177427, >99% identity, 820 bp) (Mammeri et al. 2003) in direct orientation, probably comprising a composite transposon. The GC content for the section comprising the degrading genes and the inner IS91 transposase is 62.1%, considerably higher than the average GC content of the plasmid (41.9%) or the chromosome (38.8%), suggesting a relatively recent acquisition from a different taxon.

nBLAST search of the plasmid sequence revealed that a section of the plasmid (bases 39 746 to 53 607) is 98.8% similar to plasmid pXG03-X3 from *Acinetobacter indicus* strain XG03 (CP045128.1). This region includes a series of genes responsible for benzoate catabolism (Collier et al. 1998) flanked by two ISAcsp1 transposases in opposite orientation (100% similarity, 2967 bp).

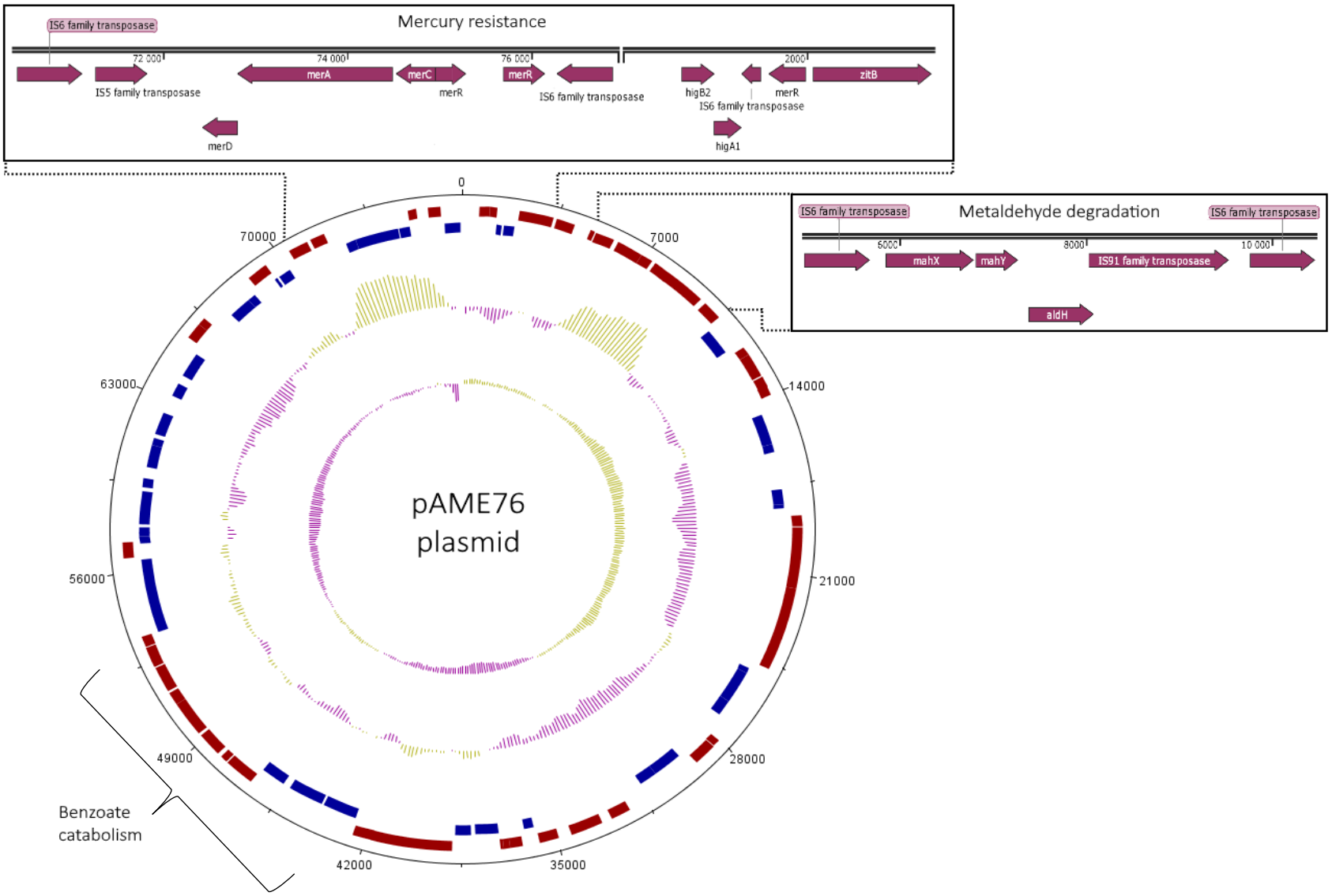


Figure 3-2. *A. calcoaceticus* E1 plasmid pAME76 map. Sequence plot was generated using Artemis and DNAPlotter (Carver et al. 2009, 2012). From outside to inside: tick marks indicate 7 000 bp steps; open reading frames for the forward and reverse strands are indicated in red and blue; GC content (window size = 1 000 bp, step size = 200 bp; purple: below average, gold: above average); innermost ring indicates GC skew (window size = 10 000 bp, step size = 200 bp; purple: below average, gold: above average).

Another section of the plasmid with higher GC content than the average includes a series of genes involved in mercuric resistance that includes *merD*, *merA*, *merC* and *merR* genes found in numerous plasmids from various *Acinetobacter* species (Mindlin et al. 2016).

3.3.1.3 Plasmid pAC10

An additional 10.5 kb circular plasmid (Figure 3-3) was found to be part of the *A. calcoaceticus* E1 genome and from here on designated pAC10. This plasmid encodes for 13 CDS including, a replication initiator protein, a relaxase, a transposase, a resolvase, two proteins involved in toxin-antitoxin plasmid stability, three hypothetical proteins with no annotation and four genes encode for other various other functions. No oriT was predicted using oriTFinder but a relaxase shares 99.07% identity with a MobA/MobL family protein from *Acinetobacter haemolyticus* (WP_161412392.1) was found, therefore this is probably a non-conjugative but mobilizable plasmid.

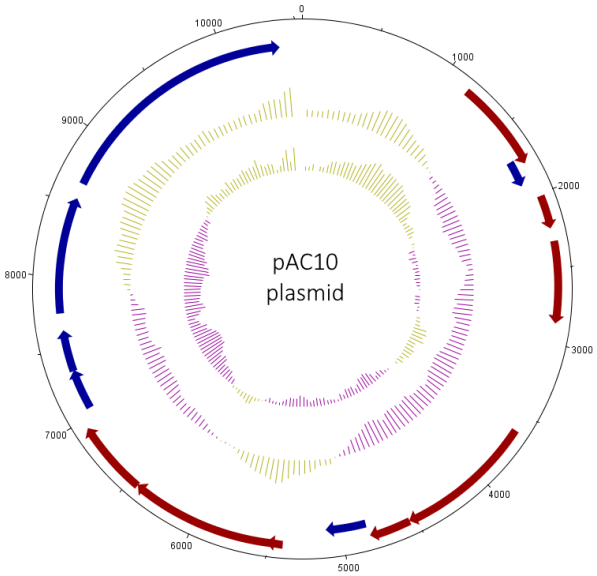


Figure 3-3. *A. calcoaceticus* E1 plasmid pAC10 map. Sequence plot was generated using Artemis and DNAPlotter (Carver et al. 2009, 2012). From outside to inside: tick marks indicate 1 000 bp steps; open reading frames for the forward and reverse strands are indicated in red and blue; GC content (window size = 1 000 bp, step size = 50 bp; purple: below average, gold: above average); innermost ring indicates GC skew (window size = 1 000 bp, step size = 50 bp; purple: below average, gold: above average).

3.3.2 Location and genetic context of the metaldehyde-degrading gene cluster in the degrading strains

The complete sequence of plasmid pAME76 from *A. calcoaceticus* E1 can be used as a reference to compare the genetic context of the metaldehyde-degrading genes in the rest of the isolates that share the same initial degradation pathway (Figure 3-4). In this strain the metaldehyde-degrading genes *mahX*, *mahY, aldH* and a partial IS91 family insertion sequence with a truncated origin and lacking the terminal section (1827-1863 and 1-111 bp missing respectively) are flanked by two intact copies of the IS*Our*1 insertion sequence. Also, two 45 bp partial repeated end fractions of a similar IS91 family insertion sequence are present upstream of the degrading genes. The sequence of the *aldH* gene seems to have been interrupted by the insertion of the IS91 family element, resulting in an open reading frame containing both the partial *aldH* gene and a fraction of the insertion sequence that extends until a stop codon is reached.

Although in the other two *Acinetobacter* strains (*A. bohemicus* JMET-C and *A. lwoffii* SMET-C) the comparison is limited by the size of the contig obtained through Illumina sequencing, it is shown that end fractions of the IS*Our*1 insertion sequences also surround the metaldehyde-degrading cluster and the IS91-like elements in these strains. In fact, the 4307 bp contig from *A. bohemicus* JMET-C is identical to the corresponding section from *A. calcoaceticus* E1. Meanwhile the 4229 bp contig from *A. lwoffii* SMET-C is 99.97% similar to these two. These two facts further suggest the horizontal transfer of the degrading cluster and the IS91 internal insertion sequence through the action of the flanking IS*Our*1 insertion sequences in the *Acinetobacter* strains.

In the case of the contig from *P. vancouverensis* SMET-B, the IS*Our*1 insertion sequences are not present either in the contig or in the whole-genome sequencing data. Also, a mutation in the *aldH* gene has generated a premature stop codon in the sequence, rendering only a 12 amino acid open reading frame for the gene.

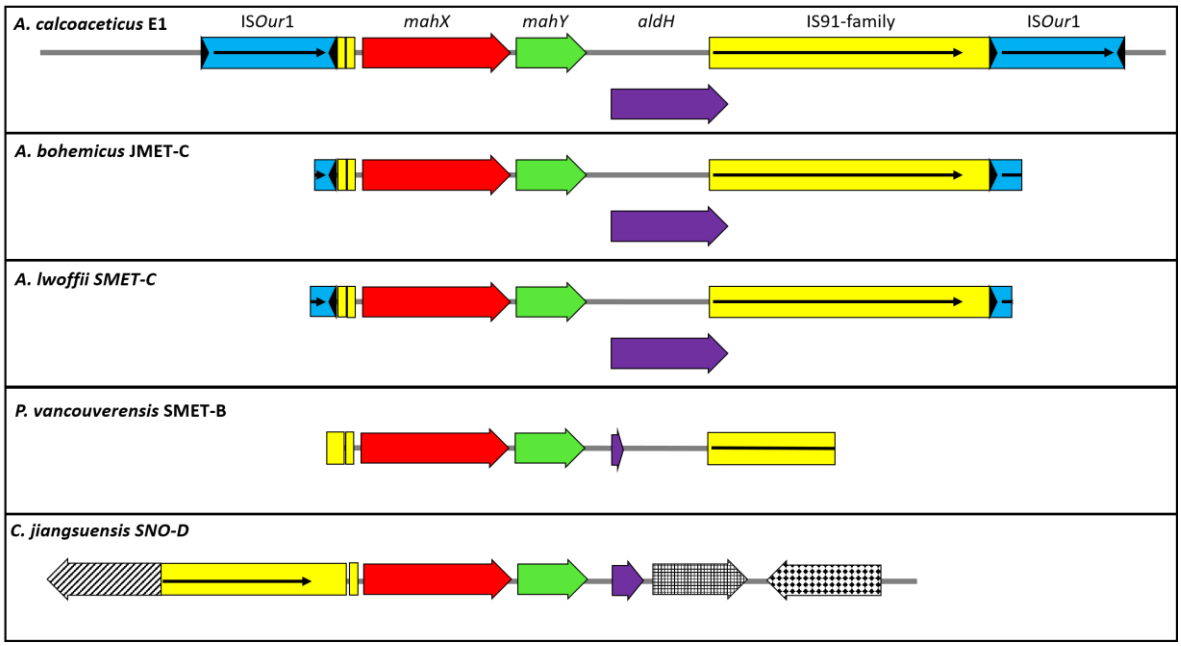


Figure 3-4. Immediate genetic context of metaldehyde degradation genes in metaldehyde-degrading strains that share the same initial pathway. Black triangles represent terminal inverted repeats. Black horizontal arrows represent transposase ORFs. Gene and insertion sequence names are indicated for *A. calcoaceticus* E1 and colour coded for the rest of the strains. For *C. jiangsuensis* SNO-D tilted line pattern indicates a partial IS21 transposase (96.7% identity; PKO63873.1); mesh pattern indicates a predicted thiolase family protein and tile pattern indicates a sequence related to sensor histidine kinase proteins. Gene maps not to scale.

In *C. jiangsuensis* SNO-D the IS*Our*1 insertion sequences are also absent in the contig as well as in the genome. A fraction of the IS91 family insertion sequence is present but upstream from the degrading gene cluster instead of downstream. In this case the *aldH* gene is also truncated by a premature stop codon generating a 99 amino acid open reading frame. Other unrelated genes are located downstream from the cluster.

To explore the possibility that the same or a similar plasmid harbours the metaldehyde-degrading operon in these other strains a search and comparison of the proteins encoded by pAME76 in the other metaldehyde-degrading isolates was carried out (Figure 3-5).

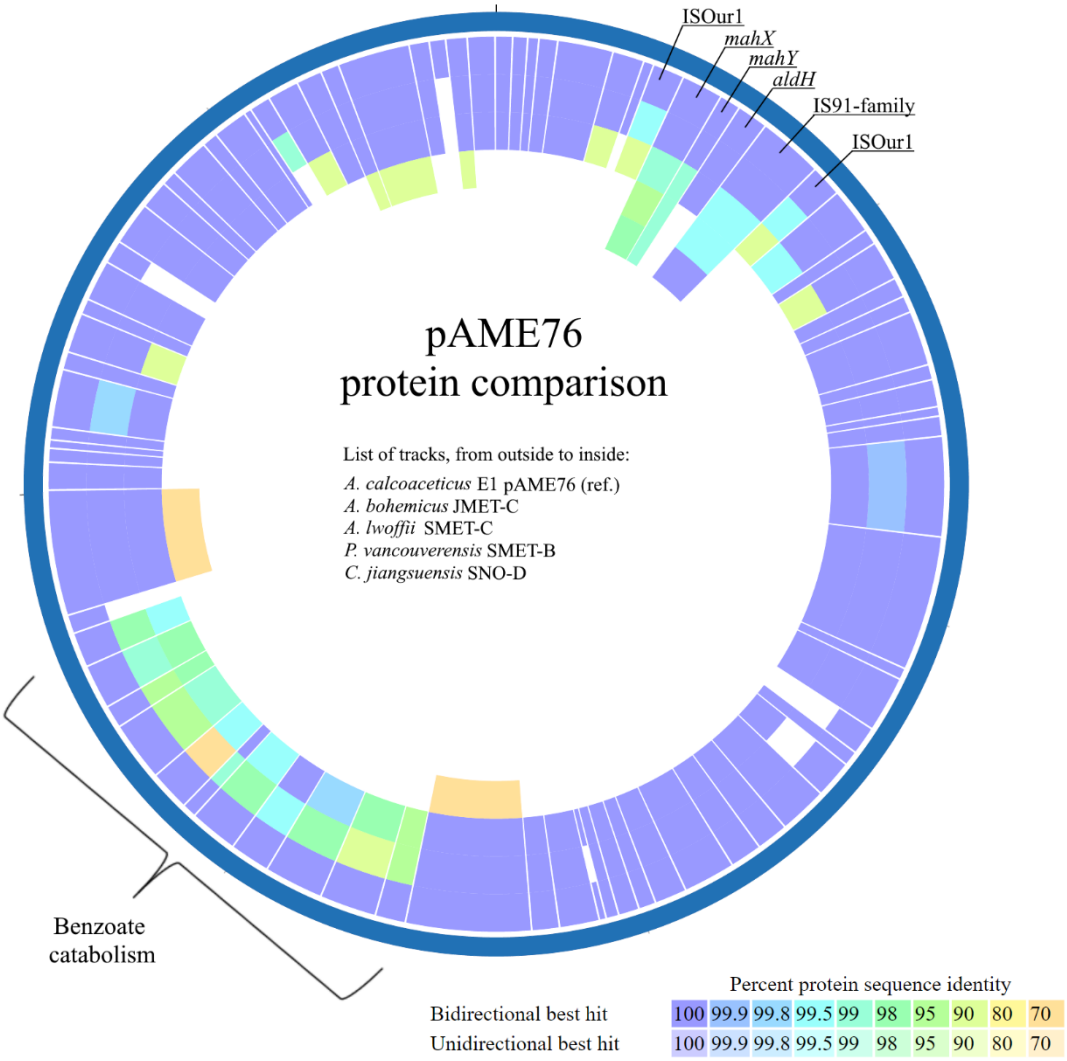


Figure 3-5. Diagram of proteins encoded by pAME76 and presence of similar sequences in metaldehyde-degrading strains that share the same degradation pathway. The analysis was performed using PATRIC Proteome Comparison Service (Wattam et al. 2017). Minimum cut-off values of 70% identity and 40% coverage were established.

Absence of most of the pAME76 plasmid proteins for degraders *P. vancouverensis* SMET-B and *C. jiangsuensis* SNO-D (innermost rings) revealed that in these isolates the metaldehyde-degrading operon is not located in a pAME76-like plasmid. On the other hand, most of the proteins in pAME76 have homologs in *A. bohemicus* JMET-C and *A. lwoffii* SMET-C, including an identical *repB* replication initiation protein in both of them, indicating that they may contain at least one plasmid with a similar replication machinery that could be harbouring the degrading operon. However, genes involved in benzoate compound catabolism, although present, generally showed a lower similarity than those of the rest of the plasmid.

To explore this observation in more detail, the reads from the whole genome sequencing of *A. bohemicus* JMET-C and *A. lwoffii* SMET-C were directly mapped against the pAME76 scaffold using Burrows-Wheeler Alignment tool (BWA) (Li and Durbin, 2009) in UGENE (Golosova et al., 2014; Okonechnikov et al., 2012). The analysis revealed two prominent drops in the read mapping coverage of the pAME76 plasmid scaffold for both isolates (Figure 3-6). The first one (26 to 28 kbp position approx.) corresponded to the absence of a *lexA* repressor in both degrading isolates, whilst the second one (42 to 53 kbp position approx.) indicated a much lower coverage of the benzoate catabolic region when compared to the rest of the plasmid. In contrast, read mapping coverage for the metaldehyde-degrading gene operon (5.8 to 8.1 kb position approx.) was relatively high.

Sequencing coverage for genes located in plasmids is usually much higher than the coverage for those located in the chromosome, owing to the usually higher copy number of plasmids in prokaryotic genomes (Bankevich et al. 2012; Providenti et al. 2006). These data suggested that in strains JMET-C and SMET-C the benzoate-degrading pathway is chromosomally encoded and, more importantly, that the metaldehyde-degrading gene operon is probably contained in plasmids in each of these two isolates, as it is in strain E1. Nonetheless, these plasmids are not identical to pAME76, although they may have several CDS in common.

To predict the putative gene composition of the plasmids present in strains JMET-C and SMET-C, PlasmidSPAdes (Antipov et al., 2016) was used with the short-read Illumina sequencing whole genome data as input. PlasmidSPAdes extracts and assembles plasmid data from whole genome sequencing projects based on sequence coverage calculations. For *A. bohemicus* JMET-C, 22 contigs (112 515 bp) were assigned to a putative plasmid containing 114 CDS including the metaldehyde-degrading operon, the mercury resistance genes and a *repB* protein identical to that of pAME76. For *A. lwoffii* SMET-C, 28 contigs (105 139 bp) were assigned to a putative plasmid containing 124 CDS also containing the metaldehyde-degrading operon, the mercury resistance genes and an identical *repB* protein. Both putative plasmids encode relaxase proteins,

which would imply a transferrable nature, and they both lack the benzoate catabolic genes. This analysis supports the idea that metaldehyde-degrading genes are also located in plasmids in strains JMET-C and SMET-C.

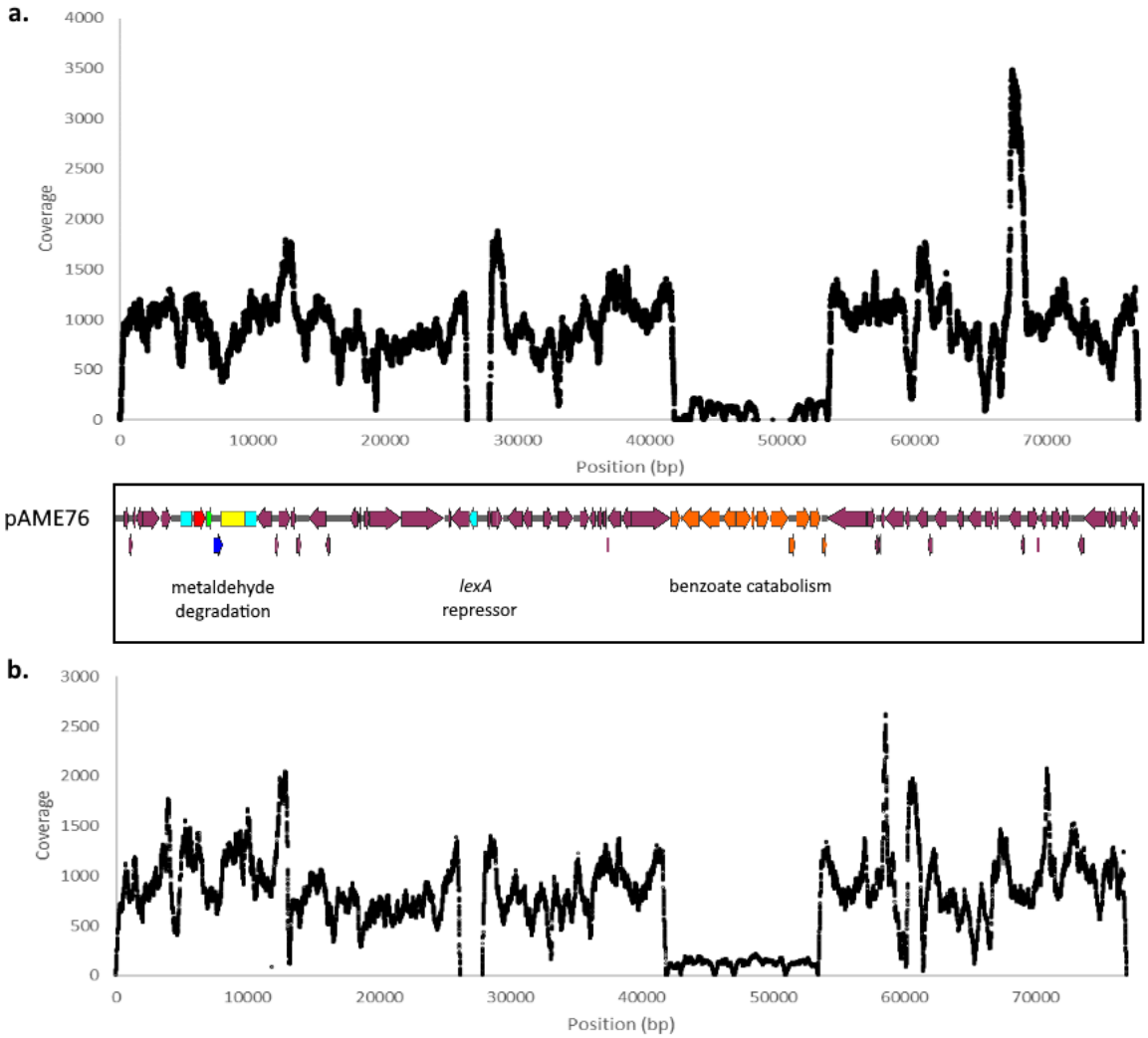


Figure 3-6. Coverage resulting of mapping the reads from **a.** *A. bohemicus* JMET-C and **b.** *A. lwoffii* SMET-C whole-genome sequencing against pAME76 plasmid reference using Burrows-Wheeler Alignment tool (BWA) (H. Li and Durbin 2009) in UGENE (Golosova et al. 2014; Okonechnikov et al. 2012).

The chromosomal or plasmid location of the metaldehyde-degrading operon in *P. vancouverensis* SMET-B and *C. jiangsuensis* SNO-D was also assessed using this approach. For strain SMET-B, 4 contigs (71 187 bp) were assigned to a putative plasmid containing 62 CDS

including the metaldehyde-degrading operon, a relaxase gene, a type IV secretion system and a type IV coupling protein, which would imply that it is a conjugative plasmid. For strain SNO-D, the metaldehyde degrading operon was not predicted to be found in a plasmid; hence, it would be either chromosomally encoded or part of a plasmid with a copy number close to one.

3.3.3 Development of a qPCR assay for quantification of metaldehyde-degrading genes

Nucleotide sequences for metaldehyde-degrading genes *mahX* and *mahY* from the isolated strains sharing the same degradation pathway were aligned using MUSCLE (Edgar 2004) (Figures 3-7 and 3-8). *aldH* genes were not taken into consideration for developing qPCR primers because of the lack of sequence conservation between all strains. The analysis revealed that for gene *mahX* 9/945 positions showed a substitution in any of the five different sequences. For gene *mahY* this number was only 1/450. The number of base substitutions per site (Table 3-8) was calculated using the Maximum Composite Likelihood model (Tamura et al. 2004). All the genes in the cluster are identical for strains *A. calcoaceticus* E1 and *A. bohemicus* JMET-C. Overall, the genes show a high degree of conservation, which facilitated primer design in shared regions.

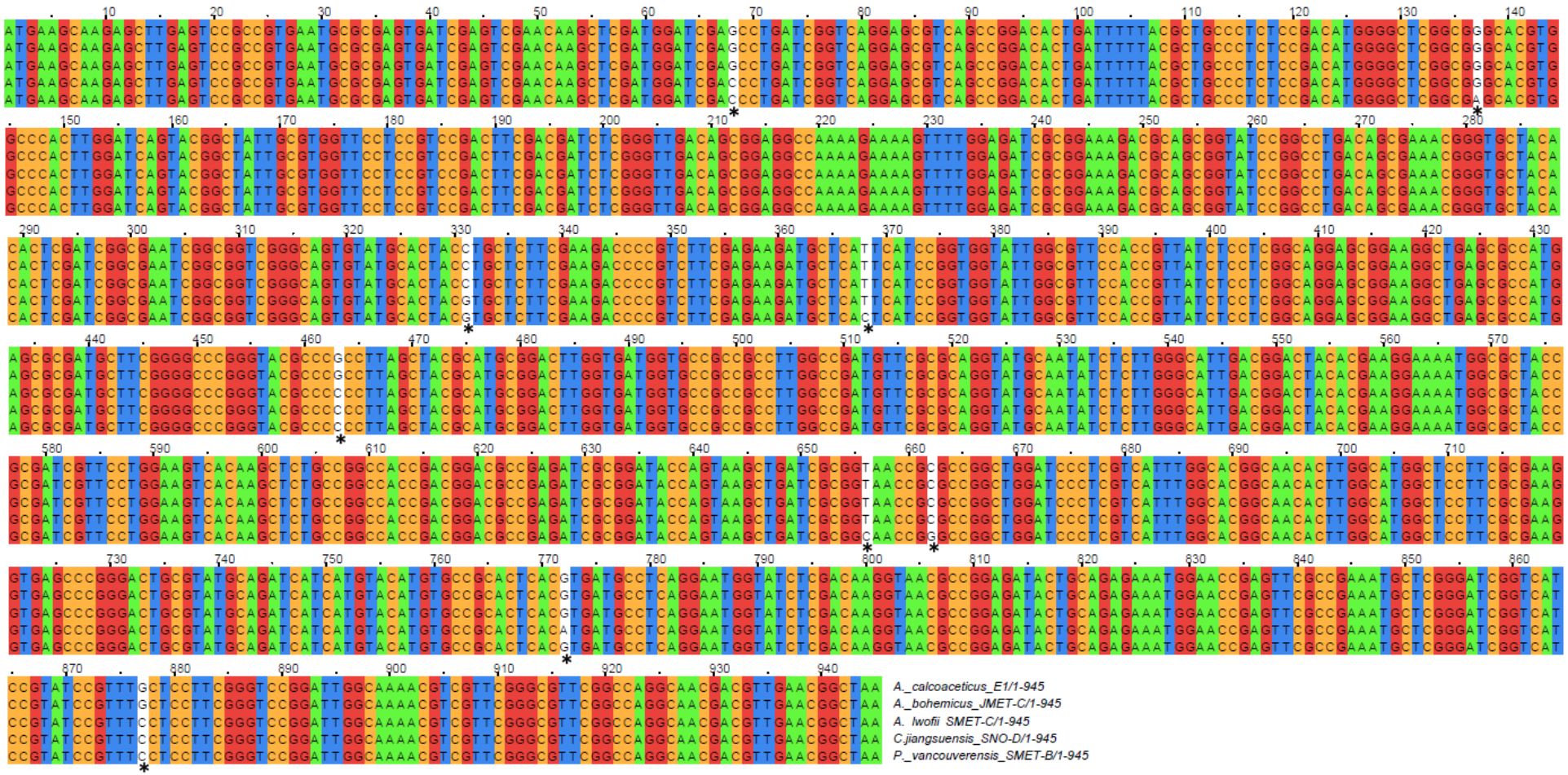


Figure 3-7. Sequence alignment for *mahX* gene from 5 different metaldehyde-degrading strains using the same degradation pathway. Asterisk indicates mismatches for at least one of the sequences.

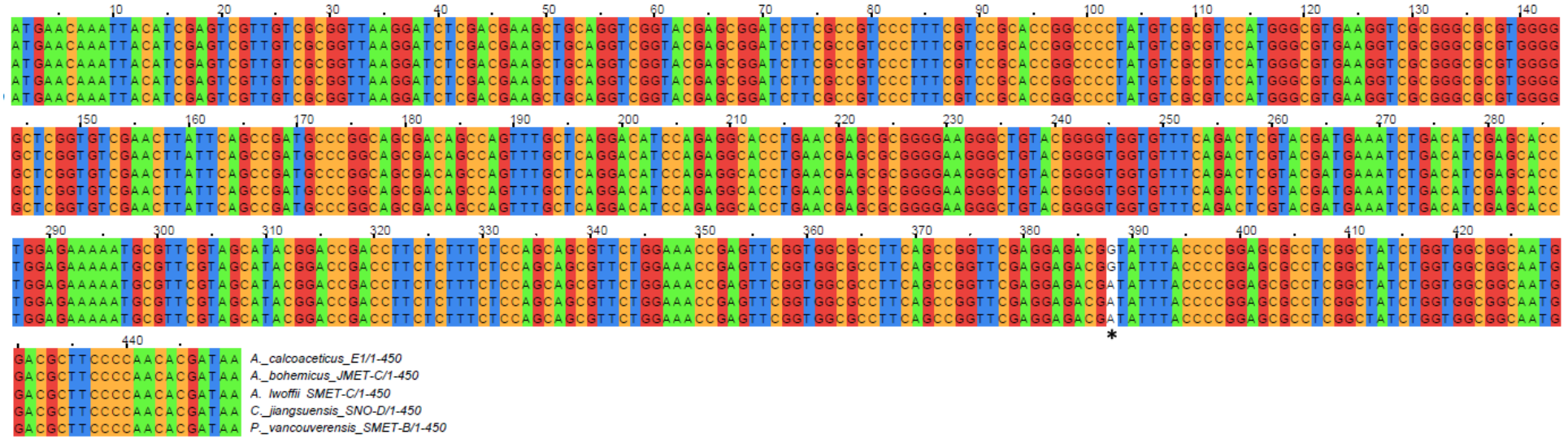


Figure 3-8. Sequence alignment for *mahY* gene from 5 different metaldehyde-degrading strains using the same degradation pathway. Asterisk indicates mismatches for at least one of the sequences.

a. ***mahX***	**E1**	**JMET-C**	**SMET-C**	**SNO-D**	**SMET-B**
***A. calcoaceticus*_E1**					
***A. bohemicus*_JMET-C**	0.0000				
***A. lwofii*_SMET-C**	0.0021	0.0021			
***C. jiangsuensis*_SNO-D**	0.0053	0.0053	0.0032		
***P. vancouverensis*_SMET-B**	0.0085	0.0085	0.0064	0.0053	

b. ***mahY***	**E1**	**JMET-C**	**SMET-C**	**SNO-D**	**SMET-B**
***A. calcoaceticus* E1**					
***A. bohemicus* JMET-C**	0.0000				
***A. lwoffii* SMET-C**	0.0022	0.0022			
***C. jiangsuensis* SNO-D**	0.0022	0.0022	0.0000		
***P. vancouverensis* SMET-B**	0.0022	0.0022	0.0000	0.0000	

Table 3-8. Estimates of evolutionary divergence between sequences. The number of base substitutions per site from between sequences are shown. Analyses were conducted using the Maximum Composite Likelihood model (Tamura et al. 2004). **a.** *mahX* gene: analysis involved 5 nucleotide sequences and there was a total of 945 positions in the final dataset; **b.** *mahY* gene: analysis involved 5 nucleotide sequences there was a total of 450 positions in the final dataset. Analyses were conducted in MEGA X (Kumar et al. 2018).

Specific primers for the metaldehyde-degrading genes *mahX* and *mahY* were designed and initially tested using endpoint PCR. Optimized PCR amplification of the genes *mahX* and *mahY* from soil DNA samples during metaldehyde degradation and their respective unexposed controls are shown in Figure 3-9 and Figure 3-10 respectively. Very similar patterns were obtained for both genes, which was expected since both genes are present in the same cluster and no significant similarity was detected with other genes though BLAST searches on primer design.

Patterns for soil SMET, in which metaldehyde had a half-life of just 3.9 d showed a relatively weak background amplification of the genes at time 0 d; nonetheless, by time 16 d the reaction was already strong, showing a slight decline in subsequent time points. On the other hand, the unexposed control soil only showed the background amplification throughout the entire assay.

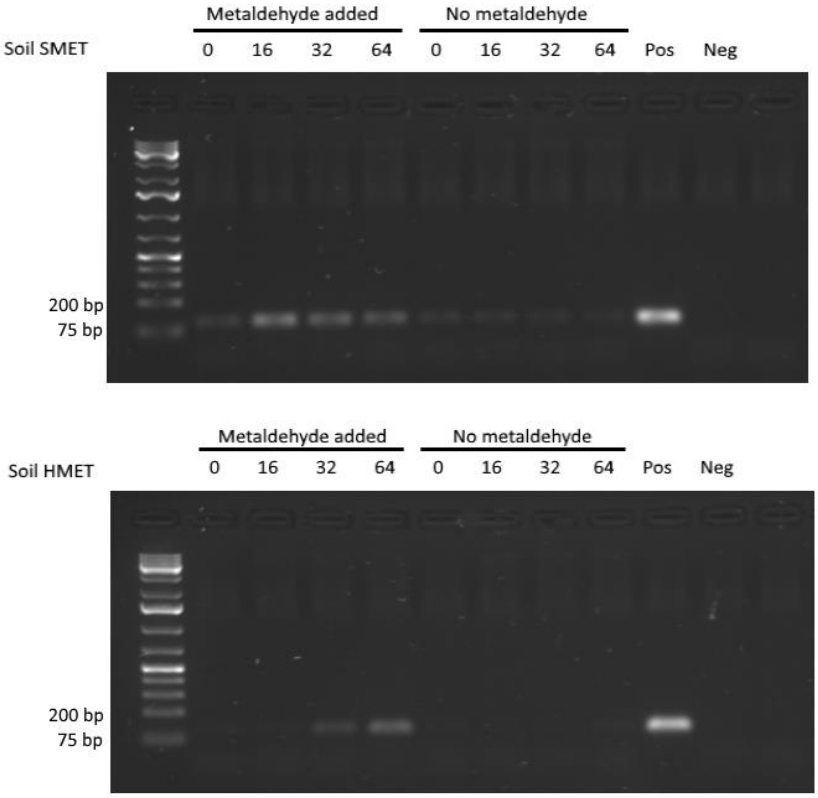


Figure 3-9. Amplification under optimized conditions using primer pair mahX-86F/183R for soil samples analysed during metaldehyde degradation studies throughout a 64 d incubation.

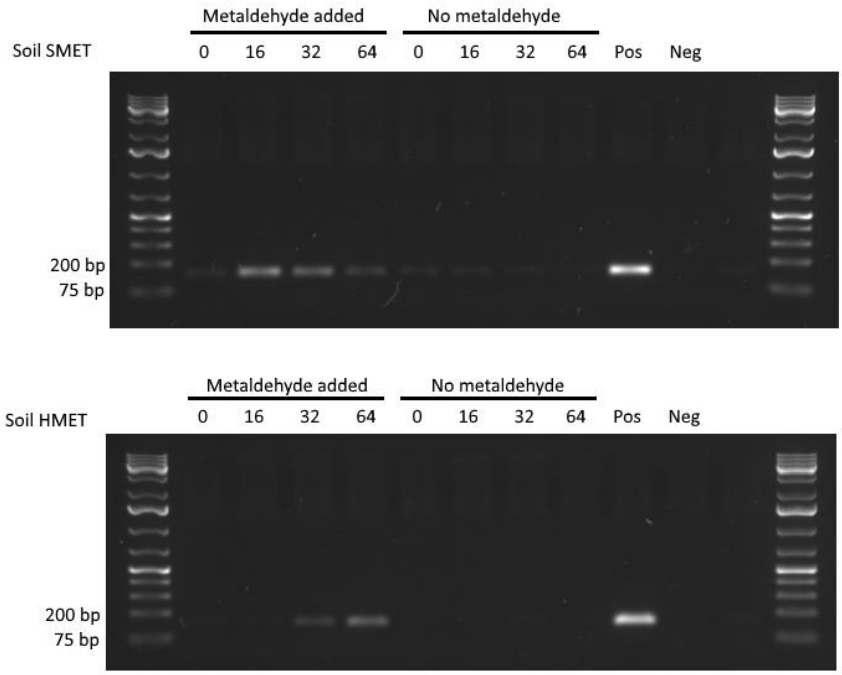


Figure 3-10. Amplification under optimized conditions using primer pair mahY-143F/287R for soil samples analysed during the first round of metaldehyde degradation studies throughout a 64 d incubation.

For soil HMET, in which metaldehyde had a much longer half-life (31.7 d), patterns did not show an initial background amplification, and it was until the 32 d time point that a positive reaction was detected. The unexposed control did not show any evident amplification. Overall, positive reactions for both soils seem to coincide with the expected increases in the number of metaldehyde degraders through time, according to metaldehyde-degradation assays indicating that the primers for genes *mahX* and *mahY* are able to provide good estimates of the metaldehyde-degrading populations.

Initial assessment of the intercalating dye qPCR method for genes *mahX* and *mahY* showed efficiencies of 66% and 77% respectively, with R^2 values of 0.99 in both cases. The method could detect the *mahX* gene with a sensitivity of 10^3 copies per reaction; on the other hand, the sensitivity for *mahY* was better, 10^2 copies per reaction. Given these results, gene *mahY* was chosen for further qPCR optimization. An increase in primer concentration from 200 nM to 500 nM resulted in an improved qPCR efficiency of 91%, R^2 value of 0.98 however the sensitivity of 10^2 gene copies per qPCR reaction did not change. The calibration curve for these conditions is presented in Figure 3-11. Melting curves from samples and standards showed single peaks and agarose gel electrophoresis after the qPCR run showed single bands of the desired size (145 bp) with no other amplicons, therefore, the assay was considered specific.

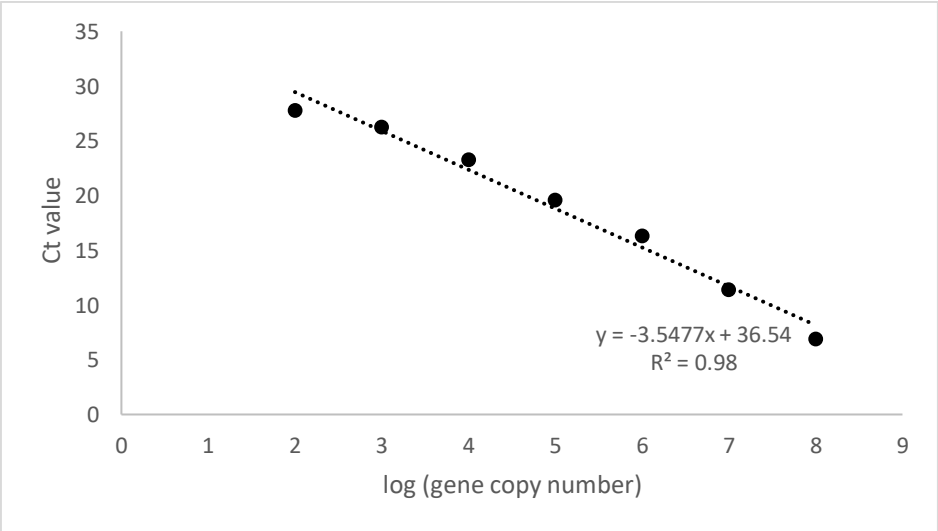


Figure 3-11. Gene copy number vs. Ct values for *mahY* gene standards at different concentrations using intercalating dye-based PCR.

Notwithstanding these encouraging results, initial tests with soil samples showed that better sensitivity was required to be able to quantify the gene copies in most soils and time points. This was related to the fact that metaldehyde degraders make up a very low percentage of the overall bacterial population in soil (Chapter 2). To increase the sensitivity of the assay, a nucleotide probe for use with the primer pair mahY-143F/287R was designed and used with the inhibitor-resistant Kapa Probe Force qPCR Master Mix (Kapa Biosystems - Roche, Cape Town) for degrading gene quantification.

The use of the fluorescent nucleotide probe coupled with the increased concentration of primers and BSA that were used were efficient in decreasing the limit of detection of the assay from 10^2 copies per reaction (12 500 copies g^{-1} soil) with the SybrGreen chemistry, to 10^1 copies per reaction (1 250 copies g^{-1} soil) with the probe-based assay. The efficiency was calculated at 96.5%, and very good linearity was obtained (r^2=0.9956, Figure 2). Given these results, gene detection using probe-based qPCR was considered sufficiently optimized.

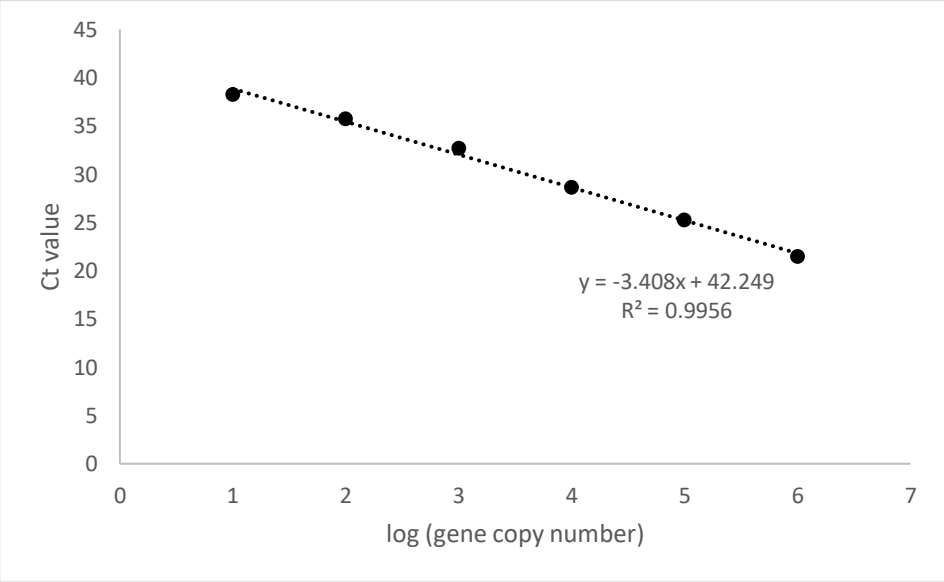


Figure 3-12. Gene copy number vs. Ct values for *mahY* gene standards at different concentrations using intercalating probe-based PCR.

3.3.4 Metaldehyde degradation profiles in soil

Physical and chemical characteristics of the soils used were determined and are shown in Table 3-9. Metaldehyde degradation profiles for STA, PA, ING and LE soils after a single application at 15 mg kg^{-1} are shown in Figure 3-13. To quantify the persistence of metaldehyde in the soil samples, data regression for the degradation profiles was performed using Single First Order or modified Hockey-stick models (Table 3-10) (FOCUS 2006). With a half-life of 0.65 d, degradation was much faster in soil STA than other soils.

Table 3-9. Physical and chemical characteristics of soils used in this study

Plot	Silt (%)	Clay (%)	Sand (%)	Soil Class	Organic Matter (%)	pH	Max WHC (% dry mass)
STA	45.35	13.88	40.77	Sandy Silt Loam	5.2	7.5	68.5
PA	51.01	13.5	35.49	Sandy Silt Loam	3.7	5.8	67.4
ING	38.63	9.37	52	Sandy Loam	3.7	6.8	63.1
LE	50.2	18.17	31.63	Clay Loam	3	7.8	61.7

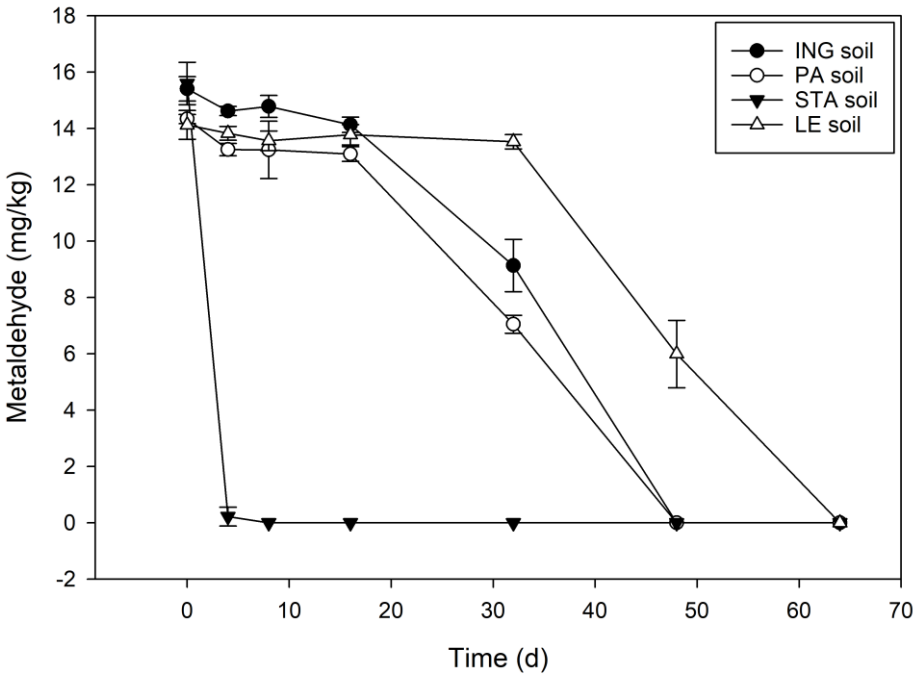


Figure 3-13. Metaldehyde degradation profiles in freshly collected soils after an initial application of 15 mg kg^{-1} soil. Bars represent standard deviation for three replicates.

Table 3-10. Regression statistics for Single First Order and modified Hockey-stick fittings of metaldehyde removal in laboratory-incubated allotment soil samples.

Soil	Storage	Metaldehyde application	Regression model	k_1 (d^{-1})	k_2 (d^{-1})	Half-life (d)	r^2
STA	Fresh	Single	SFO*	-	1.0670	0.65	1.0000
PA	Fresh	Single	MHS†	0.004890	0.6378	32.0	0.9984
ING	Fresh	Single	MHS†	0.004759	0.6992	32.3	0.9995
LE	Fresh	Single	MHS†	9.98E-004	0.5660	47.7	0.9994

* Single First Order
† Modified Hockey-stick

3.3.5 Microbial community changes in soil microcosms

3.3.5.1 Quantification of the metaldehyde-degrading gene *mahY* in soils

Quantification of *mahY* gene copies was carried out in four soils from Chapter 2 (allotment soils collected in York) and four newly collected soils from other regions of Northern England, for a total of eight different soils. Quantification was performed after a single metaldehyde application at 15 mg kg^{-1} soil. In four soils (S-Met, S-NoMet, H-Met, STA) *mahY* gene copies were successfully quantified, while for the other four soils (H-NoMet, PA, ING, LE), *mahY* gene copies were undetectable. Progression of *mahY* gene copy numbers during metaldehyde removal in the former soils and their respective controls is shown in Figure 3-14.

The graphs clearly show peaks of *mahY* gene copy numbers coinciding with moments of active metaldehyde removal in the soils. It is also evident that control soils with no metaldehyde addition showed stable lower levels of degrading gene copies. These data confirm that metaldehyde degraders are proliferating in the soil in response to metaldehyde addition. Also, soils with a higher initial number of degrading gene copies (S-Met and STA), showed faster removal (shorter half-lives) than soils with a lower initial number (H-Met and S-NoMet), which indicates that this parameter effectively provides an indication of metaldehyde degradation potential and activity in soils in which this degradation pathway is present.

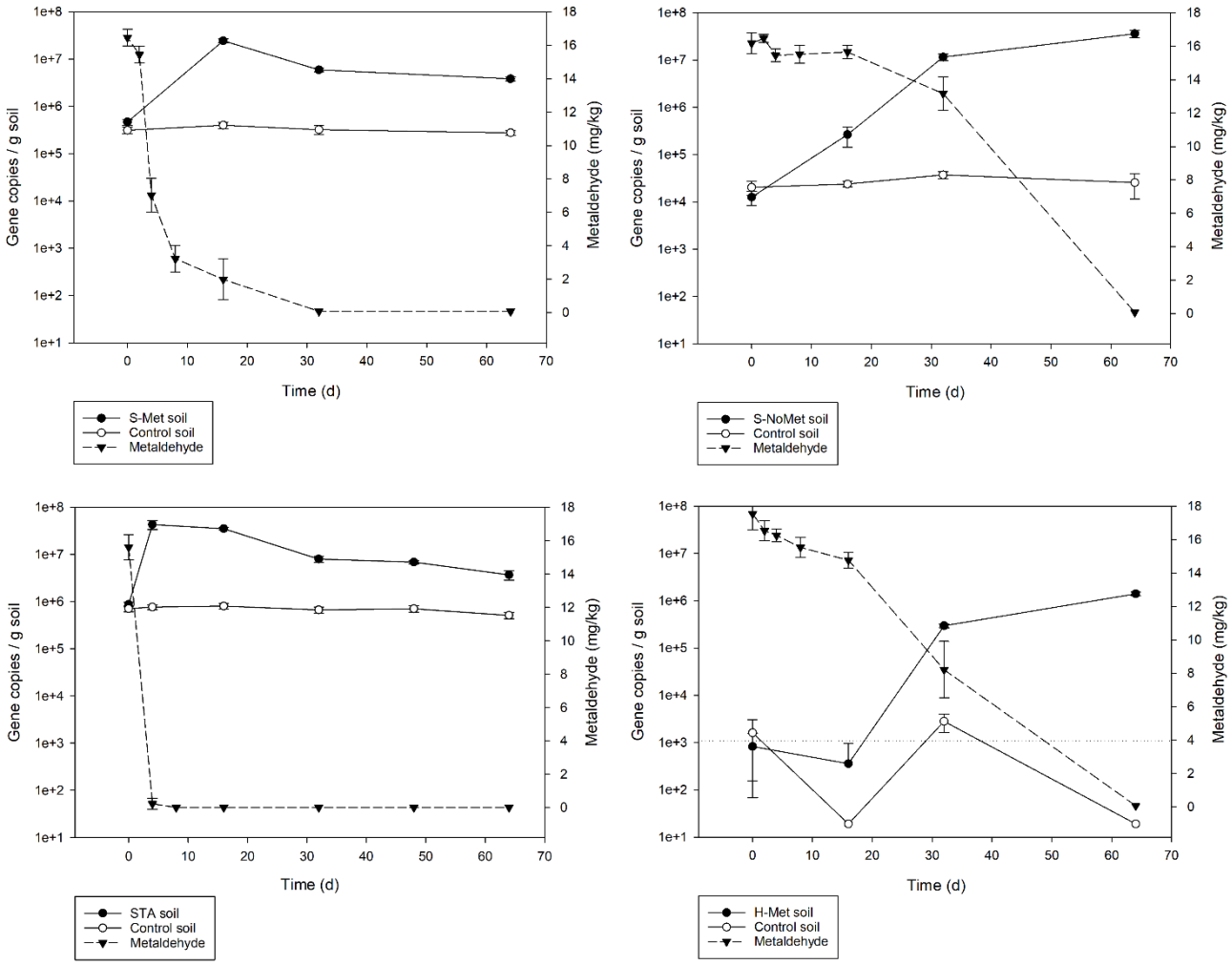


Figure 3-14. *mahY* gene copy number and metaldehyde concentrations during metaldehyde removal in different soils after an initial application of 15 mg kg^{-1} soil. Dashed horizontal line indicates the estimated sensitivity of the qPCR technique for accurate quantification of gene copies.

Nonetheless, in the rest of the soils, *mahY* gene amounts undetectable even in moments when metaldehyde removal from the soil was actively occurring (data not shown). This may be explained if the microorganisms that are effectively metabolizing the compound in these soils are using alternative degradation mechanisms and thus, cannot be detected using this assay. The fact that gene copies throughout metaldehyde removal were undetectable in soil H-NoMet, in which the only successfully isolated degrader did not harbour the described degrading genes (Chapter 2), supports this hypothesis. Inhibitory effects on PCR in the genomic DNA extracts

from these soils preventing amplification were discarded because addition of these extracts in different amounts did not repress previously positive reactions (Figure 3-15).

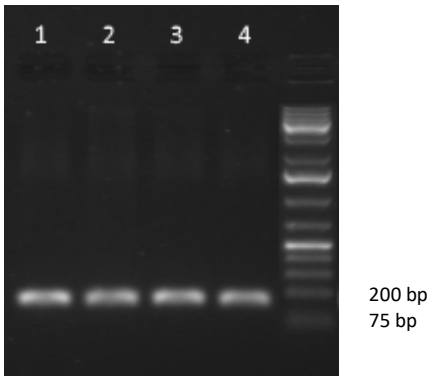


Figure 3-15. Lanes 1-4: representative example of PCR inhibition test using soil S-Met (16 d) as a positive control and adding different amounts of H-NoMet soil DNA. Lane 1: 1 µL S-Met DNA; Lane 2: 2 µL H-NoMet DNA, 1 µL S-Met DNA; Lane 3: 1 µL H-NoMet DNA, 1 µL S-Met DNA; Lane 4: 0.1 µL H-NoMet DNA, 1 µL S-Met DNA.

3.3.5.2 Changes in the microbial community assessed through 16S rRNA gene amplicon sequencing

Quality statistics for sequencing runs are shown in Annex 3. Top taxa for which abundance was increased upon metaldehyde exposure versus non exposed controls are shown in Table 3-11.

For soil STA two different unassigned species of *Acinetobacter* showed a heavily increased abundance (more than 1000-fold) when compared to the control soil not exposed to metaldehyde at time 4 d. The abundance of these species at time 0 d was 0.036 % and 0.020 %, and it increased to 0.907 % and 0.892 % respectively at time 4 d (25-fold and 45-fold increase respectively). Moreover, two strains of *Acinetobacter* were isolated as metaldehyde degraders using enrichment cultures from this soil, one of them corresponded to a strain of *A. calcoaceticus* and the other one to a strain of *A. bohemicus*, both with 16S rRNA gene sequences identical to the isolates *A. calcoaceticus* E1 (Thomas et al. 2017) and *A. bohemicus* JMET-C (Chapter 2). Furthermore, these sequences were identical to the partial Illumina read sequences obtained for the most enriched taxa from the soil.

Table 3-11. Identification of top ten taxa for each soil for which abundance is increased in samples exposed to metaldehyde in the laboratory versus non-exposed controls upon metaldehyde removal below the limit of detection*

Soil	Family	Genus	Species	Fold increase
STA	Moraxellaceae	*Acinetobacter*	Unassigned	1382
	Moraxellaceae	*Acinetobacter*	Unassigned	1359
	Sphingomonadaceae	*Sphingobium*	Unassigned	118
	Planococcaceae	*Sporosarcina*	Unassigned	90
	Enterobacteriaceae	Unassigned	Unassigned	58
	Streptomycetaceae	*Streptomyces*	*reticuliscabiei*	56
	Unassigned[1]	Unassigned	Unassigned	55
	Paenibacillaceae	Unassigned	Unassigned	55
	Unassigned[2]	Unassigned	Unassigned	52
	Solirubrobacteraceae	*Solirubrobacter*	Unassigned	47
PA	Sphingomonadaceae	*Sphingomonas*	*wittichii*	905
	Pseudomonadaceae	*Pseudomonas*	*umsongensis*	370
	Acidobacteriaceae	Unassigned	Unassigned	304
	Xanthomonadaceae	Unassigned	Unassigned	286
	Sphingobacteriaceae	*Mucilaginibacter*	*kameinonensis*	242
	Comamonadaceae	*Rubrivivax*	*gelatinosus*	216
	Chitinophagaceae	*Niabella*	Unassigned	198
	Unassigned[3]	Unassigned	Unassigned	147
	Unassigned[4]	Unassigned	Unassigned	128
	Sphingobacteriaceae	*Pedobacter*	Unassigned	114
ING	Thermoactinomycetaceae	Unassigned	Unassigned	115
	Rhodobacteraceae	Unassigned	Unassigned	98
	Acetobacteraceae	*Roseomonas*	Unassigned	86
	Unassigned[5]	Unassigned	Unassigned	84
	Lachnospiraceae	*Coprococcus*	Unassigned	83
	Microbacteriaceae	Unassigned	Unassigned	80
	Oxalobacteraceae	*Duganella*	nigrescens	76
	Chitinophagaceae	*Niabella*	Unassigned	67
	Rhizobiaceae	*Rhizobium*	Unassigned	66
	Paenibacillaceae	Unassigned	Unassigned	65
LE	Saprospiraceae	*Haliscomenobacter*	Unassigned	874
	Comamonadaceae	*Hydrogenophaga*	Unassigned	528
	Moraxellaceae	*Acinetobacter*	*rhizosphaerae*	478
	Unassigned[6]	Unassigned	Unassigned	409
	Unassigned[7]	Unassigned	Unassigned	342
	Unassigned[8]	Unassigned	Unassigned	331
	Cytophagaceae	*Siphonobacter*	*aquaeclarae*	327
	Rhodocyclaceae	*Zoogloea*	Unassigned	322
	Sphingomonadaceae	*Sphingomonas*	*wittichii*	275
	Comamonadaceae	*Rhodoferax*	Unassigned	262

*Analyses were performed with samples from time points 4 d, 48 d, 48 d and 64 d for soils STA, PA, ING and LE respectively.

[1]This taxon could only be identified to the order level: Actinomycetales. [2]This taxon could only be identified to the order level: Sva0725. [3]This taxon could only be identified to the class level: Gemm-2. [4]This taxon could only be identified to the class level: Elusimicrobia. [5]This taxon could only be identified to the class level: SHA-26. [6]This taxon could only be identified to the class level: TSBW08. [7]This taxon could only be identified to the class level: ML635J-21. [8]This taxon could only be identified to the class level: Betaproteobacteria.

Samples from STA, PA, ING, and LE soils were rarefied to 153 442, 135 185, 134 171 and 98 638 reads respectively.

Remarkably, for soil PA, the most enriched taxa corresponded to species of *Sphingomonas wittichii* and *Pseudomonas umsongensis*, which could potentially correspond to

metaldehyde degraders, since closely related taxa were isolated as degraders in Chapter 2. Nevertheless, since enough isolates had already been obtained from the other soils previously, no selective enrichment for degraders was attempted in this case. For soil LE, potential degraders on the list were again *Sphingomonas wittichii* and *Acinetobacter rhizosphaerae*. For soil ING, no easily identifiable potential degraders were found to be listed.

3.4 Discussion

3.4.1 Generation of reference-quality whole genome sequence of metaldehyde degrader *A. calcoaceticus* E1 and genetic context of the degrading genes

Prokaryotic genome sequencing has become increasingly affordable in recent years mainly because the increasing availability and decreasing price of sequencing technologies such as Illumina short-read sequencing. This technology is capable of generating very accurate reads (>99%) with a length ranging from 75 to 300 bp (or around 150–600 bp if pair-end sequencing is used) (Goodwin et al. 2016). Typically, a bacterial genome sequenced using Illumina Mi-Seq short read sequencing only will result in tens or hundreds of contigs for every genome. For instance, previous short-read sequencing for *A. calcoaceticus* E1 (Chapter 2) resulted in 98 contigs, with the metaldehyde-degrading operon present in a single contig with a 4229 bp length and consisting of only four genes.

Generation of longer contigs is mainly limited by the presence of repeated regions that are longer than the read length used, because they cannot be disambiguated and are thus often fragmented into contigs (S. Goldstein et al. 2019) or incorrectly collapsed into a single sequence in the case of paralogs (Ekblom and Wolf 2014). These repetitive regions include ribosomal genes, transposons, insertion sequences, CRISPR arrays, amongst others (Kingsford et al. 2010).

Insertion sequences and transposable elements are particularly abundant in plasmids when compared to their host chromosome (Siguier et al. 2014). These sequences allow accessory genes from plasmids to mobilize under evolutionary pressure, but they also inhibit

plasmid sequence assembly and therefore limit the amount of information that can be gathered on their genetic context (George et al. 2017).

Sequencing technologies such as the ones developed by Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT) can overcome these problems by using long reads to resolve complex genomic regions. Even though they suffer from high error rates (~ 13% for PacBio, ~5% for ONT) (Goodwin et al. 2016; Oxford Nanopore Technologies 2020), these long reads can be used as scaffolds for the contigs generated by the highly accurate short-read data (Ashton et al. 2015) or they can be used to disambiguate regions of the assembly graph (Wick et al. 2017), being able to produce reference-quality genomes in either case. This last approach was used for the strain E1, fully resolving the chromosome and plasmid composition of the strain, the gene copy number, the plasmid location of the metaldehyde-degrading gene cluster and its position within a putative composite transposon.

Xenobiotic-degrading genes for an extensive list of compounds have been found to be located in mobile genetic elements (MGE), such as transferrable plasmids, integrative and conjugative elements (ICEs), and transposons (Dennis 2005; Shintani and Nojiri 2013; Top et al. 2002; Van Der Meer and Sentchilo 2003). These elements have in common that they are all able to move within and/or between genomes, thus allowing “evolution in quantum leaps” through horizontal gene transfer (HGT) (Hacker and Carniel 2001). Presence of nearly identical pathways and enzymes for xenobiotic degradation, shared by phylogenetically unrelated bacterial strains are thought to mainly be a consequence of this process (Nojiri et al. 2004).

The plasmid responsible for metaldehyde degradation in strain E1, pAME76 (76.9 kbp), shares with many other catabolic plasmids the characteristic of being a relatively large plasmid (>50 kb) (Nojiri et al. 2004).

Not all plasmids have the same ability to be transferred to different cells. Conjugative plasmids should encode for a type IV secretion system, which forms the mating pore, a type IV coupling protein, a relaxase and have an origin of transfer (oriT) region (Cabezón et al. 2015). If

a plasmid lacks the type IV secretion system and coupling protein it can potentially still be transferred to another cell by hijacking the type IV secretion system encoded by another MGE within the same cell, being a mobilizable or transferrable plasmid (Smillie et al. 2010).

Classically, if a plasmid lacked the relaxase, it would have been characterized as a non-mobilizable or non-transferrable plasmid. Studies have shown that these type of plasmids corresponded to 48% of those detected in Proteobacteria (Smillie et al. 2010). Plasmid pAME76 would fall into this group, lacking detectable genes for mating pair formation, type IV coupling protein and relaxase, as well as a readily identifiable oriT region. However, recently it has been demonstrated that certain "non-mobilizable" plasmids can be mobilized to a different host by using the relaxase encoded by another plasmid within the same cell through the "relaxase in trans" mechanism (Moran and Hall 2019; Ramsay and Firth 2017), nonetheless this has been demonstrated only with small plasmids so far.

Even if pAME76 is a genuinely non-mobilizable plasmid, it harbours numerous insertion sequences and transposons (24 complete or truncated transposase gene copies), including the two IS*Our*1 insertion sequences flanking the metaldehyde degrading genes and the IS91-like elements, possibly comprising a composite transposon. Catabolic genes are often flanked by insertion sequences, which may have facilitated its acquisition by a specific replicon, but may also ease the subsequent exchange of the genes between different replicons and hosts (Springael et al. 1996; Wyndham et al. 1994). This transposon-like configuration is also found in the other degrading *Acinetobacter* strains, JMET-C and SMET-C, which suggests that it may have mediated the horizontal transfer of the operon between species in the genus.

IS*Our*1 belongs to the IS6/IS26 family comprising 65 insertion sequences. The full sequence of IS6 has never been fully determined but it is thought to be identical or nearly identical to IS26, which has been the subject of most of the experimental studies in the family (Harmer and Hall 2019). IS26 has a very important role in the generation of genetic diversity in Gram-negative bacteria, and particularly relating to antibiotic resistance in *Acinetobacter*

(Blackwell et al. 2016; Post and Hall 2009). IS*Our*1 itself has been implicated in the acquisition of cephalosporin resistance genes from *Acinetobacter* into the chromosome of *Oligella urethralis* (Mammeri et al. 2003). Both IS26 and IS*Our*1 are 820 bp long, with 14 bp inverted repeats and a single gene which encodes a 234 amino acid transposase (Mollet et al. 1985). IS*Our*1 is also very similar (>99% nucleotide similarity) to IS1008 (AJ251307) and according to specific criteria regarding insertion sequence classification, may be considered the same insertion sequence (Harmer and Hall 2019).

The mechanism of transposition of IS*Our*1 specifically has not been studied experimentally, however the details of this process in the flagship of the group, IS26, have been elucidated and other members of the family seem to share the same mechanism, making it likely that all members are able to use the same transposition routes (Harmer and Hall 2020).

IS26 movement is different from most of the other families of insertion sequences and has been shown to be mediated by a transposase via two main routes: a copy-in route and a targeted conservative route. A graphical representation of these two routes can be found in Annex 4 (Harmer et al. 2014). The copy-in route involves two molecules, a donor molecule with a copy of IS26 and a target molecule which are covalently bound to form a cointegrate, a molecule formed when joining both. Here the insertion sequence and the random 8 bp target site are duplicated. The resulting molecule would have two copies of IS26 in direct orientation at the boundaries between the original molecules. The cointegrate can in some cases be resolved by homologous recombination between the two directly oriented copies of the insertion sequence in a recombination proficient host, which would render a single IS26 sequence surrounded by the 8 bp target site duplication in the target molecule, resembling a product of a simple transposition (Iida et al. 1984; Mollet et al. 1983, 1985). This mechanism is exclusive of IS26 family members (Siguier et al. 2015). In the targeted conservative route both molecules carry a copy of IS26. It is targeted because it occurs at one or other end of the two insertion sequences instead of a random site. It is conservative because the insertion sequence

and the target site are not duplicated, no bases are lost or gained (Harmer et al. 2014; Harmer and Hall 2016). Its efficiency if 50-fold higher than the copy in mechanism when both molecules have a copy of IS26 (Harmer et al. 2014).

Regarding the IS91-like elements, sequence analysis has shown that the majority of IS91 family isoforms do not appear in the classic way other insertion sequences do, as direct repeats flanking adjacent open reading frames forming composite transposons (Garcillán-Barcia and de la Cruz 2002; Mahillon and Chandler 1998). Instead it has been shown that transposition of adjacent genes can occur through a "one-ended" mechanism involving a single insertion sequence, which is translocated along with adjacent sequences of different lengths (Garcillán-Barcia and de la Cruz 2002). IS91 uses a rolling circle mechanism for transposition (Garcillán-Barcia et al. 2001) which in principle is initiated at the *ori* end of the insertion sequence (3' to the transposase, where transposition starts) and terminates at the *ter* end (5' to the transposase, where transposition ends). However, termination can fail in around 1-10% of the events, and transposition is extended to neighbouring sequences, until a surrogate end sequence is reached (Garcillán-Barcia et al. 2002; Tavakoli et al. 2000); this can also happen if the *ter* region is deleted (Mendiola et al. 1994) Related insertion sequences presumably capture neighbouring genes in the same way (Toleman et al. 2006).

Given that in the *Acinetobacter* strains and in the *P. vancouverensis* SMET-B strain the *ter* region of the IS91-like insertion sequence is absent, one might speculate that the neighbouring genes, the metaldehyde-degrading gene cluster, could have been transposed along with the IS91-like insertion sequence in first instance and afterwards the IS*Our*1-mediated transposition might have taken place for the *Acinetobacter* strains only, as suggested by the fact that the IS91-like sequence is interrupted by an intact IS*Our*1 sequence and not the other way around. Also, it has been suggested that IS91 family elements might be hot spots for insertion of other transposable elements (Garcillán-Barcia and de la Cruz 2002). Nonetheless, the fact that IS91 family elements do not generate target site duplications upon insertion (Diaz-Aroca et al.

1987) and IS26 insertion sequences may or may not do it depending on the specific transposition mechanism they use (Harmer et al. 2014) it is difficult to trace the specific evolutionary history of the cluster. Even so, one might speculate on the possible succession of events to generate the typical cluster observed in the isolated *Acinetobacter* strains (Figure 3-16).

The location of the metaldehyde-degrading gene cluster in a plasmid was corroboratedg by coupled long-read and short read sequencing in *A. calcoaceticus* E1, and *in silico* plasmid reconstruction using PlasmidSPAdes indicates this is also the case for 3 of the 4 remaining strains. Of course, the bioinformatics analysis of plasmid reconstruction from short-read sequencing has limitations. It has been shown that this approach has a precision of approximately 0.75, which means that a quarter of the predicted plasmid contigs might be false positives, and also, multiple plasmids from a single genome can be merged into a single bin (Arredondo-Alonso et al. 2017). Nevertheless, obtaining the same result repeatedly in different strains and certainty that the sequence is undoubtedly plasmid-based in *A. calcoaceticus* E1 indicates that a false positive for plasmid location is highly unlikely in these cases.

Even though the plasmids in all the metaldehyde degrading strains that use this pathway are not identical, horizontal gene transfer of the degrading cluster has undoubtedly occurred. The presence of the IS91 and IS26 family elements in the cluster and the different GC nucleotide percentages between the gene cluster (59.81-62.74%) and the whole genome sequences of metaldehyde-degrading *Acinetobacter* strains (38.7-40.2%) are parametric indicators of horizontal gene transfer (Ravenhall et al., 2015). Moreover, the high conservation of the degrading gene cluster and its proteins between strains when compared to the lower conservation of the rest of the genome (Chapter 2) is in itself an indicator of horizontal gene transfer and even provided the means to identify the degrading genes in the first place.

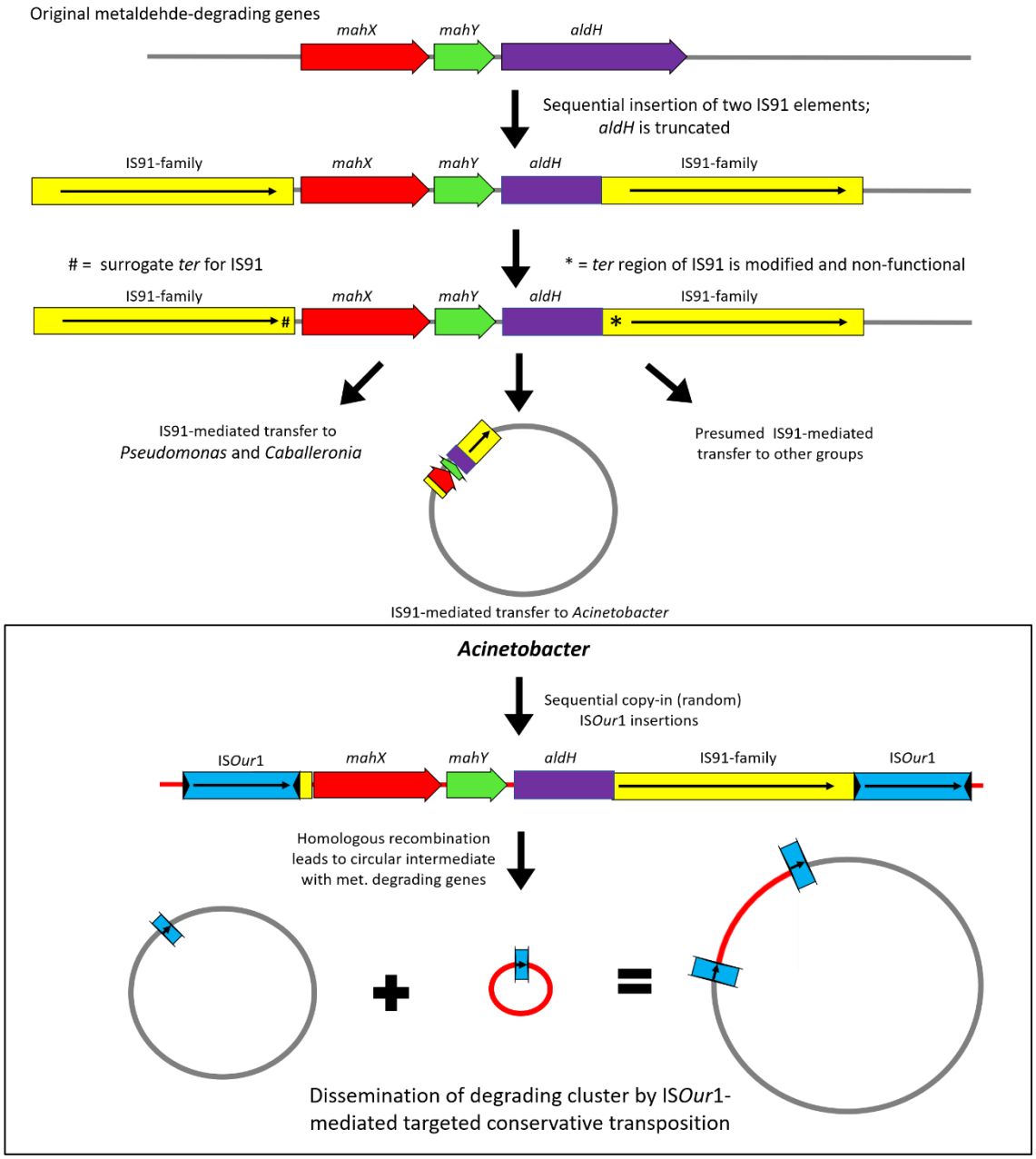


Figure 3-16. Possible evolutionary history leading to the metaldehyde-degrading gene cluster present in the isolated *Acinetobacter* strains.

3.4.2 Development of a qPCR assay for quantification of metaldehyde-degrading genes

MPN assays were used in the previous chapter to be able to quantify metaldehyde degraders in soil samples, however the identification of a set of genes involved in metaldehyde degradation made the development of a molecular assay to quantify the degrading population possible. A qPCR test was designed to be able to accurately detect the degrading genes in the cluster. Genes *mahX* and *mahY* were chosen as the most promising candidates. Several reasons

led to the use of *aldH* gene to be discarded for this purpose. First, aldehyde dehydrogenases are widely distributed in all domains of life; for instance, in an analysis of 258 strains of the genus *Pseudomonas* there was a median of 24 ALDH-coding genes per strain (Riveros-Rosas et al. 2019), which poses a serious challenge to the specificity of the primers to be designed for use in environmental samples. Also, the *aldH*-encoded protein showed a high similarity to other proteins present in the NCBI database (up to 78%), so there is an increased chance of cross-reaction. In fact, two different sets of primers that had been initially designed for the *aldH* gene either did not generate strong bands on the gels or produced unspecific amplifications (additional bands of different molecular weight, data not shown) which could not be eliminated by increasing the annealing temperature.

Second, in all the strains the full sequence of the *aldH* gene from the cluster is either interrupted by the IS91-family insertion sequence or truncated by a premature stop codon which would generate loss of active site and substrate binding residues (Joernvall and Persson 2006; Riveros-Rosas et al. 2013). This suggests that the acetaldehyde generated upon metaldehyde degradation is probably converted to acetate by a different ALDH than the one present in the gene cluster. This agrees with the observations by Thomas (2016) who found that a chromosomally encoded 55.3 kDa ALDH was enriched 34-fold when *A. calcoaceticus* E1 was grown in metaldehyde as the sole carbon source versus when grown on acetate. Since the activity of the specific *aldH* gene in the cluster would not be essential for metaldehyde metabolism, it would be of little use for detecting degrading potential when better candidate genes are available.

Genes *mahX* and *mahY* showed very similar patterns in the initial endpoint PCR assays in different soils, however, upon qPCR optimization, *mahY* had a better performance in terms of qPCR efficiency and sensitivity, therefore it was chosen as the ideal target for qPCR assays. We were able to show that for at least half of the tested soils the *mahY* assay could effectively detect the presence of metaldehyde degraders and its proliferation upon metaldehyde addition;

furthermore, when a high number of initial gene copies was detected (>10^5 copies g^{-1} soil) metaldehyde half-lives were very short (0.65-2.6 d). These data show that *mahY* gene copies quantified by qPCR effectively provide an indication of metaldehyde degradation potential and activity when the degradation pathway is present in the soil.

Many studies have found a correlation between degrading gene copy number and pesticide degradation/mineralization activity (Bælum et al. 2012; Rousidou et al. 2017; Sagarkar et al. 2013), highlighting the value of this approach for biodegradation monitoring. Pre-exposed soils usually harboured higher degrading gene copies than non-exposed soils (Lal *et al.*, 2015; Castillo *et al.*, 2016). However, it is important to highlight the limitation that soils with similar genetic degrading potential can, in some instances, exhibit different pesticide-degrading activities, depending on factors such as soil type (Martin-Laurent et al. 2004) and pH (Yale et al. 2017). This emphasizes the value of assessing biodegradation potential/activity using several different strategies simultaneously.

On the other hand, half of the soils did not show gene copy numbers above the limit of detection, even upon metaldehyde addition and degradation. Very low or undetectable numbers of degrading genes have been observed in instances where elimination was negligible (Nousiainen et al. 2014), others in which degradation was mainly through abiotic processes (Ben Salem et al. 2018), or when degradation was presumably occurring through alternative biological pathways for pesticides such as diuron (Pesce *et al.*, 2013).

In our case metaldehyde elimination invariably occurred in these soils and modelling of metaldehyde degradation kinetics showed a marked lag phase before the onset of fast elimination (modified hockey-stick model), a hallmark of biological degradation (FOCUS 2006). Therefore, the most likely explanation is that metaldehyde removal was occurring through an alternative biological pathway in these soils. The fact that other strains using a different metaldehyde degrading pathway were isolated in this study (Chapter 2) and in a previous study (Thomas et al. 2017) supports this idea. Elucidation of these alternative pathways will provide

new opportunities for development of complementary tests for metaldehyde degradation potential in soils.

Regarding the increased abundance of specific taxa in metaldehyde-exposed soils vs. non-exposed controls it is important to point out that it is indeed possible to use these data to identify metaldehyde degraders, as in the case of soil STA, in which the most enriched taxa corresponded with degrading strains, being corroborated by selective enrichment procedures. This indicates that, even though not in every case, the relative abundance data may aid in the identification and subsequent isolation of xenobiotic degraders.

3.5 Conclusions

In this chapter it is shown that this newly described metaldehyde-degradation pathway is often harboured in plasmids in degrading bacterial strains and was probably dispersed through neighbouring insertion sequences and transposons throughout different taxa. Furthermore, the development of a sensitive molecular assay for degrading gene detection was described and the technique was successfully tested and found to be useful in a variety of soils. A highly informative follow up study could potentially involve the use of the radiolabelled pesticide to be able to correlate degrading gene copy numbers with mineralization rates. Also, upon elucidation of alternative metaldehyde degradation pathways that undoubtedly exist in nature, the development of complementary assays for degrading gene detection in soil could be pursued.

CHAPTER 4: REMOVAL OF ENVIRONMENTALLY RELEVANT CONCENTRATIONS OF THE PESTICIDE METALDEHYDE FROM WATER IN BIOAUGMENTED PILOT-SCALE SLOW SAND FILTERS[4]

4.1 Introduction

The European Union Drinking Water Directive (Council Directive 98/83/EC; DWD) defines the regulatory limits for individual and total pesticides in drinking water at 0.1 µg L^{-1} and 0.5 µg L^{-1} respectively. More than 90% of the failures to comply with these limits in England in 2014 and 2015 were caused by metaldehyde (Chief Inspector of Drinking Water 2017a), and even though the total number of failures has been reduced in recent years, mainly through catchment management, metaldehyde remains the single biggest challenge to the supply of pesticide free water in the UK (Chief Inspector of Drinking Water 2018; Mohamad Ibrahim et al. 2019).

It has proven problematic for water companies to attain cost-effective removal of metaldehyde using conventional drinking water treatment processes based on adsorption to activated carbon (Castle et al. 2017; Dillon et al. 2013) and while more sophisticated treatments such as Advanced Oxidation Processes (AOPs) are effective for metaldehyde clean-up (Autin, Hart, et al. 2013b), they are extremely costly to implement.

[4] The work presented in Chapter 4 was possible due to a collaboration with researchers from Cranfield University (Laura Pickering (PhD student) and Francis Hassard (Lecturer in Public Health Microbiology). A three-month placement at Cranfield University was undertaken. Cranfield University provided metaldehyde reference material, equipment, reagents and culture media for growth and inoculation of strains, equipment and reagents for flow cytometry, sand and raw water for all assays. Metaldehyde quantification by LC/MS-MS was performed at Cranfield University with equipment and technical support of their technical staff. Installation and operation of slow sand filters, and physicochemical analyses of influent and effluent water were carried out by Cranfield University's researchers. The original experimental design of this chapter was my own work and was slightly modified with input from Cranfield University's researchers for improvement and practical considerations. The chapter is being prepared for peer-review and publication. The data presented in this chapter are subject to a client confidentiality agreement. Technical performance characterisation of the slow sand filters will feature in a Cranfield University doctoral thesis authored by Laura Pickering.

In this context, biological strategies such as slow sand filtration emerge as a promising option for metaldehyde removal. Slow sand filters have been used for water purification since the 19th century and they are advantageous compared to other treatment strategies because of their simplicity, low chemical and energy requirements, and high level of water treatment (Haig et al. 2011). In conventional slow sand filters, autochthonous microbial populations from the source water attach to the sand and colonize the surfaces while utilizing nutrients present in the water for growth (Benner et al. 2013).

Slow sand filters can be efficient in removing organic micro-pollutants and have been proposed as viable options for advanced treatment in small wastewater treatment plants (Escolà Casas and Bester 2015). Biodegradation and sorption have been shown to be the most relevant mechanisms for organic contaminant elimination in this type of systems (Scheytt et al. 2004). Nevertheless, only certain operational biofilters have been capable of removing metaldehyde.

Even though under certain conditions the microbial community in sand filters might adapt to degrade specific micropollutants, this usually happens after continuous fluxes of the compound for lengthy periods of time (Zearley and Summers 2012) and/or with exposure to much higher concentrations than those that would usually be present in the water to be purified. For instance, (Rolph et al. 2019) could successfully achieve compliant water after acclimation of the microbial community in sand by exposure with 50 µg L^{-1} of metaldehyde for 5 d. Their data supported the hypothesis of an increase in slow-growth microorganisms able to use metaldehyde directly as a primary carbon source. Nonetheless, using this strategy at larger scale would involve adding the contaminant in relatively high concentrations to side stream reactors for a relatively long period of time, with an associated risk of contaminant breakthrough being involved. These acclimation experiments also suggested that only a small group of specialist degraders might be responsible for its uptake and removal in biofilters.

Metaldehyde-degrading bacteria have recently been isolated for the first time and proven to use the compound as a sole source of carbon (Thomas et al. 2017). Since then, the

diversity of bacteria known to be capable of metaldehyde degradation has been expanded and a shared degradation pathway has been identified (Chapter 2).

The indigenous microbial communities in biologically-active sand filters seldom contain dominating populations capable of mineralizing specific micropollutants that can be present in drinking water sources at the ng L^{-1} to µg L^{-1} range (Benner et al. 2013). Furthermore, ensuring extended contact times between the microbial community and the contaminant by setting longer hydraulic retention times (as used in slow sand filters) can lead to more efficient micropollutant removal (Zearley and Summers 2012; Zuehlke et al. 2007). This is the context in which bioaugmentation of slow sand filters with suitable specialist degrading organisms emerges as a logical next step. However, most candidate strains for bioaugmentation procedures have only been shown to utilize their specific substrate at relatively high concentrations (much higher than micropollutant levels), and it is in most cases unknown if they can actually degrade it when immersed in an environmental matrix at very low concentrations (Benner et al. 2013).

Therefore, the aims of this chapter were: i) to identify the most suitable bioaugmentation agents from a library of metaldehyde-degrading strains through bench-scale assays; ii) to evaluate if bioaugmentation of upscaled systems with selected strains is an effective strategy for removing the contaminant, and iii) to determine if the presence of an active population of metaldehyde degraders is the key factor governing its removal.

4.2 Materials and Methods

4.2.1 Chemicals and reagents

Metaldehyde (99%) was purchased from was purchased from Acros Organics (Morris, NJ), methanol (≥99.8%%) was purchased from Fisher Scientific (Loughborough, UK), all other chemicals were purchased from Sigma-Aldrich (St. Louis, MO.),

4.2.2 Analytical methods for metaldehyde

Analyses of metaldehyde concentrations were performed from filtered aqueous samples (0.22 µm) by LC-MS/MS using ultra high-performance liquid chromatography (ExionLC AD, AB Sciex, Framingham, MA) coupled to a triple quadrupole mass spectrometer (QTRAP 6500 plus, AB Sciex). Chromatographic separation was done at 60°C through an ACQUITY UPLC BEH C18 1.7 µm 2.1 x 50 mm and a Vanguard BEH C18 1.7 µm precolumn by injecting 10.0 µL samples. 2 mM ammonium formate in water (A) and in methanol (B) were used as mobile phases. The mobile phase flow was 0.45 mL min^{-1} starting with 5% of component B and gradually changing to 95% B over 2 min, which was kept constant for 0.40 min. A 0.10 min gradient back to 5% B followed and then 1.50 min at initial conditions. ESI capillary voltage of 5500 V and desolvation temperature of 200°C were used. During chromatography, metaldehyde is given a charge via an $NH4^+$ adduct encountered in the mobile phase and so the precursor ion is selected at the mass of metaldehyde + ammonium: m/z = 194.1. The product ions are observed at m/z = 106 and m/z = 62 using multiple reaction monitoring with a dwell time of 150 ms. Fragmentation is achieved using a collision energy of 7 and 9 V respectively.

Calibration curves were constructed by analysing the standards at a range between 0.01 and 20 µg L^{-1} in ultrapure water. Linear regression analysis of calibration data was performed obtaining coefficients of determination ≥ 0.99. Dilutions of a certified reference material (metaldehyde, 992 µg mL^{-1}(±35 µg mL^{-1}) SPEX Organics) in ultrapure water were used to prepare 2.0 µg mL^{-1} standards which were analysed every 10 samples during the analytical runs.

4.2.3 Bacterial strains

Metaldehyde-degrading bacterial strains *Acinetobacter calcoaceticus* E1 (Thomas et al. 2017), *Acinetobacter bohemicus* JMET-C, *Acinetobacter lwoffii* SMET-C, *Pseudomonas vancouverensis* SMET-B, *Caballeronia jiangsuensis* SNO-D and *Sphingobium* sp. CMET-H were isolated and characterized previously (Chapter 2).

4.2.4 Calibration curves for OD600nm vs number of bacterial cells

Calibration curves for OD_{600nm} vs number of cells for all metaldehyde-degrading strains grown in LB broth were constructed. Each metaldehyde-degrading strain was grown in LB broth at 150 rpm, 30°C to an OD_{600nm} between 1.0 – 1.5 and diluted in LB broth to OD_{600nm} values of approximately 1.0, 0.8, 0.6, 0.4, 0.2, 0.1, and 0.05. Each suspension was further diluted 1:10 000 in sterile 0.85% NaCl solution before quantification by flow cytometry. Total bacterial cell concentration in each dilution was determined in triplicate after SybrGreen I staining using an Accuri C6 flow cytometer using fast mode with 100 µL and 50 000 events as limits.

4.2.5 Laboratory-scale batch assays for metaldehyde degradation

4.2.5.1 Batch experiments in buffer solution using pure cultures

For all bench-scale assays, assimilable organic carbon (AOC) free glassware was prepared beforehand according to a previously described method d (APHA-AWWA-WEF 2012). Removal of metaldehyde at an environmentally relevant starting concentrations of 2.0 µg L^{-1} (Rolph et al. 2019) was first assayed in pure culture for six different metaldehyde-degrading strains, in defined sterile media (filter sterilized PBS with added trace elements solution (Vishniac and Santer 1957)). The inocula were prepared by growing each strain for 20 h at 30°C and 150 rpm (32 h for *Sphingobium* CMET-H) in 200 mL of LB broth. Cells were then pelleted down (4000 g, 10 min), washed twice with PBS, resuspended in supplemented PBS to an appropriate OD, and 2.5 mL of inoculum were added to triplicate 250 mL Erlenmeyer flasks per

strain to obtain a microbial load of 1.0 x 10^7 cells mL^{-1} in 125 mL final volume. Duplicate abiotic (inoculum replaced with PBS) controls were included in the analysis. Duplicate non-degrading strain controls (*A. calcoaceticus* RUH2202) were also prepared to evaluate removal due to sorption/absorption to/by biomass. Samples were taken at selected time points, filtered (0.22 µm) and stored in chromatography vials at -20°C until metaldehyde analysis by LC-MS/MS.

4.2.5.2 Batch experiments in bioaugmented raw water/sand systems

Removal of metaldehyde at an environmentally relevant starting concentration of 2 µg L^{-1} was then assayed for the strains that showed the best performance in the previous assay, but using raw water as media and either clean sand (washed with deionized water and autoclaved) or sand with a biofilm (taken from a full scale slow sand filter with an active biofilm operated by Thames Water at Ashford Common). The inocula were prepared as described in the previous section except washes and resuspension were performed using sterile 0.85% NaCl solution instead of PBS. Triplicate 250 mL Erlenmeyer flasks per strain with a liquid volume of 100 mL, 33 g of either clean sand or sand with a biofilm and an added microbial load of 4.0 x 10^7 degrading cells mL^{-1} were prepared. Duplicate abiotic controls were also included in the analysis. Samples were taken at selected time points, filtered (0.22 µm) and stored in chromatography vials at -20°C until metaldehyde analysis by LC-MS/MS.

4.2.6 Slow sand filter assays

4.2.6.1 Construction and operation of the slow sand filter pilot plant

Six Perspex columns were constructed with at total height of 1.29 m, with an internal diameter of 0.15 m (22.8 L total internal volume each) and with horizontal sampling taps every 20 cm (GiCAM, UK). The bed depth of the sand was 0.80 m. Quartz Sand (Specialist Aggregates, UK) was sorted to a grain size of 0.1-0.3 mm and a uniformity coefficient of 1.35. The sand was

thoroughly washed and placed on a 0.20 m gravel bed with a particle size of 1-5 mm. Given the total diameter of the sand bed, wall effects were considered minimal.

A total of 25,000 L of sub-potable grade drinking water were abstracted from a Water Treatment Works in the South of England and stored in the dark for no more than 1 month prior to use at the National Research Facility in Water and Wastewater Treatment. The water was obtained from a surface water abstraction from the River Thames. The water pre-treatment consisted of reservoirs storage, coagulation-flocculation using alum, and roughing filtration prior to our sampling. The water was spiked to a final concentration of metaldehyde of 2.0 µg L^{-1}. Water was supplied to each column using a peristaltic pump (530 U, Waterson Marlow, UK) at a flowrate of 0.1 m h^{-1} (flow rate for one column: 1.8 L h^{-1}). The empty bed contact time (EBCT) was 9.5 h and the hydraulic retention time was 3.5 h. A Schmutzdecke was permitted to form and the filters were allowed to ripen for a period of 2 months of continual monitoring prior to bioaugmentation inoculation. The DOC of the source water was 4.5 mg L^{-1} with a turbidity in the inlet which was 0.1-4.5 NTU.

The slow sand filter pilots were monitored weekly for water quality parameters including total organic carbon (TOC) SM 5310 and turbidity SM 2130 (APHA-AWWA-WEF 2012), which were within the expected thresholds for operational slow sand filters (< 1 NTU and 15-25% removal respectively) (Collins 1998).

4.2.6.2 Slow sand filter treatments

Pilot-scale slow sand filters were operated for a total of 72 d. Table 4-1 shows the description of the treatments in each of the six scale slow sand filters for the first 55 days of the study (Phase 1). Figure 4-1 shows the operational pilot-scale slow sand filters.

Table 4-1. Treatments for the pilot scale slow sand filters carried out during the first 55 d of the study (Phase 1)

Sand filter number	Metaldehyde in influent	Bioaugmentation	Purpose
1	No	No	Non-treated control
2	No	Yes*	Persistence of the bioaugmentation agent without metaldehyde input
3	Yes	Yes*	Effect of bioaugmentation on metaldehyde removal
4	Yes	Yes*	Effect of bioaugmentation on metaldehyde removal
5	Yes	No	Removal of metaldehyde without bioaugmentation
6	Yes	No	Removal of metaldehyde without bioaugmentation

*Bioaugmentation with *A. calcoaceticus E1* at 1x concentration (16 d) and 2x concentration (42 d) was performed.

Figure 4-1. Pilot-scale slow sand filters: each filter had a total volume of 22.8 L and a calculated water volume of 6.325 L. Sampling ports were located at different depths throughout the filter.

From time point 56 d onwards (Phase 2), in an effort to further enhance metaldehyde removal from the filters, the strategy was modified to increase the amount of *A. calcoaceticus* E1 in a single filter (filter 3) and include *Sphingobium* CMET-H as a bioaugmentation agent as well (filters 4 and 5). Details are described in Table 4-2.

Table 4-2. Treatments for the pilot scale slow sand filters carried out from time 56 to 72 d of the study (Phase 2)

Sand filter number	Metaldehyde in influent	Bioaugmentation	Purpose
1	No	No	Non-treated control
2	No	No[1]	Persistence of the original bioaugmentation agent (*A. calcoaceticus* E1) without metaldehyde input
3	Yes	Yes[1]	Effect of bioaugmentation with an increased amount *of A. calcoaceticus* E1 (3x) on metaldehyde removal
4	Yes	Yes[1]	Effect of bioaugmentation with *Sphingobium* CMET-H on metaldehyde removal
5	Yes	Yes[2]	Effect of bioaugmentation with *Sphingobium* CMET-H on metaldehyde removal
6	Yes	No	Removal of metaldehyde without bioaugmentation

[1]Bioaugmentation had been performed previously at times 16 and 42 d with *A. calcoaceticus* E1.
[2]No previous bioaugmentation had been performed on this filter before

4.2.6.3 Inoculation of bacterial strains

A summary of the amount of inoculum in each sand filter, bioaugmentation microorganisms and inoculation times is presented in Table 4-3.

Table 4-3. Bioaugmentation treatments for the pilot-scale slow sand filters used throughout the study.

Inoculated microorganism	Time (d)	Phase	Inoculum Cn. in aqueous phase of sand filter (CFU mL^{-1})*	Inoculated filters
A. calcoaceticus E1	16	1	2.40×10^7 (1x)	2, 3, 4
A. calcoaceticus E1	42	1	4.84×10^7 (2x)	2, 3, 4
A. calcoaceticus E1	56	2	8.11×10^7 (3x)	3
Sphingobium CMET-H	56	2	5.00×10^7	4, 5

*The concentration was determined by dilution plate counts of the inoculum in LB agar and the estimated water volume in the sand filters.

For inoculation at time point 16 d (1x), *A. calcoaceticus* E1 was grown to late exponential phase (OD 600 nm = 1.5) by inoculating it in 300 mL of LB broth in 10x 500 mL Erlenmeyer flasks and incubating for 24-32 h at 30°C and 150 rpm. Cells in 375 mL of culture were pelleted down in replicate 750 mL centrifuge bottles (3200 g, 45 min), washed with 200 mL of 0.85% NaCl solution, pelleted down again (3200 g, 30 min) and finally resuspended in 75 mL of 0.85% NaCl. Resulting volumes were combined and duplicate plate counts in LB agar were performed. At time point 16 d, 200 mL of this suspension were added to slow sand filters 2 , 3 and 4. For this, the water level was lowered to 2 cm above sand bed, the inoculum was added, the water level was lowered further to sand bed level, and then raised again to the normal operating level. Addition of 0.85% NaCl instead of inoculum was performed in sand filters 1, 5 and 6 as control. For inoculation at time point 42 d (2x), the whole process was repeated but with double the amount of *A. calcoaceticus* E1 cells.

For inoculation at time point 56 (3x) d the inoculum was prepared in the same way as the first time, but the whole amount of the *A. calcoaceticus* E1 inoculum was pooled and added to sand filter 3 only, to obtain roughly triple the amount of bioaugmentation microorganisms of the original inoculum. In filters 1, 2 and 6 an equal volume of sterile 0.85% NaCl solution was added.

At this same time point, an identical inoculum preparation procedure as in the 2x addition was followed but using *Sphingobium* CMET-H grown for 36 h instead of *A. calcoaceticus* E1 as the inoculating organism, and the inoculum was added to filters 4 and 5 specifically. The former had seen two previous inoculations with *A. calcoaceticus* E1 (1x and 2x), while the latter had not been bioaugmented before.

4.2.6.4 DNA extraction from sand filter material

Sand samples were withdrawn at specific time points from the upper layer of each slow sand filter (Schmutzdecke) using a sterile spatula. Genomic DNA was extracted from 0.4 g of

sand using the NucleoSpin Soil DNA extraction kit (Macherey-Nagel, Düren, Germany). A relative 16S rRNA gene copy number quantification technique (Smets et al. 2016) was carried out by spiking the sand samples before extraction with an internal DNA standard (*Thermus thermophilus* DSMZ 46338) at an estimated 0.1% of total DNA. The quality of nucleic acid extractions was verified in SybrSafe-stained (Thermo Fisher Scientific) 1.0% agarose gels. DNA concentration and purity were quantified using a Jenway Genova Nano spectrophotometer (Cole Parmer, St. Neots, UK).

4.2.6.5 16S rRNA gene amplicon community analyses

The V4-V5 region of the 16S rRNA gene was amplified for each sand sample using 515F/806R primers (Walters et al. 2016). PCR conditions and purification steps were described elsewhere (Castro-Gutiérrez et al. 2018). 16S rRNA amplicons were sequenced on an Illumina MiSeq instrument using 300-bp paired-end sequencing at the Bioscience Technology Facility of the University of York.

Whole-community 16S rRNA V4-V5 region amplicon sequence information for the sand samples was quality checked using Fastqc software v. 0.11.9. Read sequences were analysed using QIIME2 v2020.6.0 (Caporaso et al. 2010). Demultiplexed paired-end sequences were imported and the Divisive Amplicon Denoising Algorithm 2 was implemented for quality filtering (Q score ≥ 30), chimera removal, and feature table construction (Callahan et al. 2016). Taxonomy was assigned to the feature table using the Greengenes 13.8 reference database (McDonald et al. 2012). Feature tables were rarefied to equal number of reads, and the associated taxonomy was then extracted for further analysis.

4.2.6.6 Statistical analyses

All statistical analyses were performed in PRIMER7 (Primer-E Ltd., Auckland, New Zealand). 16S rRNA gene abundance data for phylum and genus levels was standardized by

sample total, transformed by square or fourth root and Bray-Curtis resemblance matrices were constructed. Principal coordinates analysis (PCO) was used for data ordination. Permutational MANOVA (PERMANOVA) was used assess the influence of different factors in the microbial community composition (9999 permutations). PERMDISP was used to test for homogeneity of multivariate dispersions. The significance of the bacterial community clustering was also estimated using analysis of similarity (ANOSIM) in specific cases (9999 permutations). SIMPER analyses were performed to obtain the main variables (taxa) that explain the dissimilarity between samples grouped in different levels of the factors. The DIVERSE function was used to analyse the richness and diversity (Shannon's H' index), calculated from the abundance for each distinct genus in the samples.

4.2.7 Accession numbers

Raw reads for 16S rRNA amplicons were deposited in the European Nucleotide Archive under study PRJEB40595.

4.3 Results

4.3.1 Calibration curves for OD_{600nm} vs number of bacterial cells

The calibration curves for OD_{600nm} vs number of bacterial cells for metaldehyde-degrading strains are shown in Figure 4-2. All the tested strains were in the order of 10^8 cells mL^{-1} at an OD_{600nm} equal to 1.0, and good linearity was obtained as shown by R^2 values > 0.99 in all cases. These data were used to precisely calculate the dosing of metaldehyde degraders in batch scale assays.

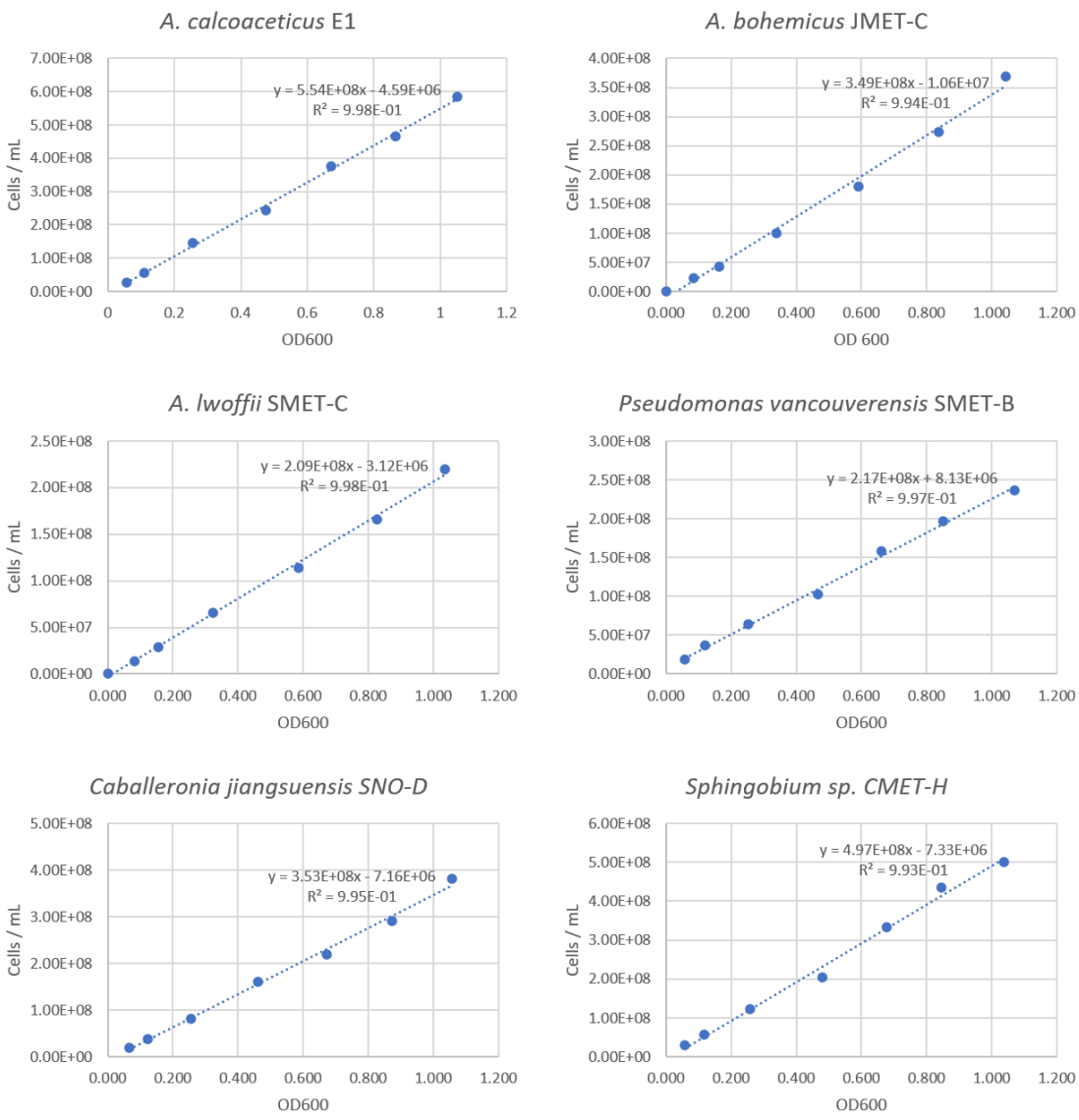


Figure 4-2. Calibration curves for OD_{600nm} vs. number of cells for metaldehyde-degrading strains grown in LB broth.

4.3.2 Removal of metaldehyde in laboratory-scale batch experiments

First it was necessary to determine if the previously isolated metaldehyde-degrading strains could remove the compound at environmentally relevant initial concentrations (2 µg L^{-1}) under ideal conditions (pure culture, defined medium and metaldehyde as the only carbon source) at a practical inoculum level (1 x 10^7 cells mL^{-1}) and how they compared to one another in terms of removal efficiency. Metaldehyde removal assays under these conditions are presented in Figure 4-3.

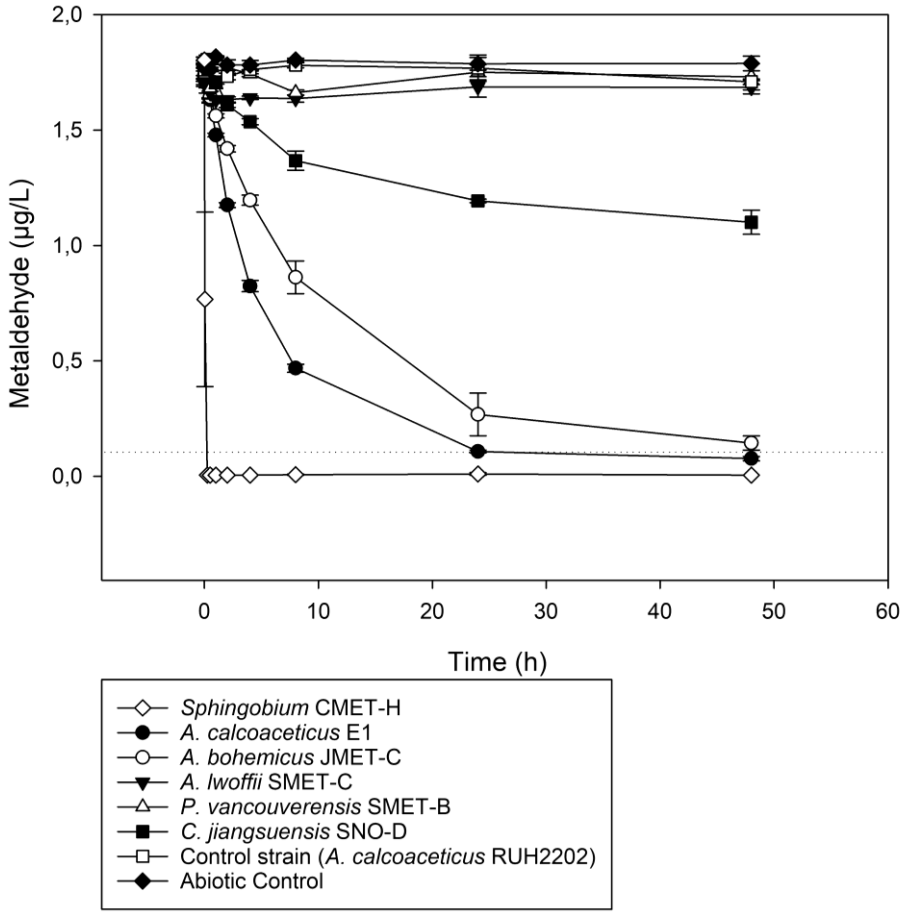


Figure 4-3. Metaldehyde removal assays for degrading strains in pure culture (1 x 10^7 cells mL^{-1}) and supplemented PBS media at an environmentally relevant starting concentration (2 µg L^{-1}). The dashed line indicates the regulatory limit (0.1 µg L^{-1}).

The different strains varied in their ability to remove metaldehyde at this low initial concentration. *Sphingobium* CMET-H showed the fastest removal, with undetectable amounts of the compound at the 2 min time point. Two of the *Acinetobacter* strains (*A. calcoaceticus* E1 and *A. bohemicus* JMET-C) were next in terms of removal efficiency. *C. jiangsuensis* SNO-D showed a slower utilization of the compound, while negligible removal was observed for *A. lwoffii* SMET-C and *P. vancouverensis* SMET-B. Marked differences were present even though most these strains (except *Sphingobium* CMET-H) share the same metaldehyde-degrading gene cluster (Chapters 2 and 3). Lack of metaldehyde removal by the control strain (*A. calcoaceticus* RUH2202) shows that elimination of the compound is not due to adsorption or absorption by the added microbial biomass.

The best performing strains, *A. calcoaceticus* E1, *A. bohemicus* JMET-C and *Sphingobium* CMET-H, were chosen for further assays in more challenging conditions more closely resembling a slow sand filter environment. These conditions included the use of raw non-sterile water and presence of sand with and without a microbial biofilm. These conditions would expose the inocula to different sources of carbon and interactions with the microorganisms in the water and sand. It was desirable that metaldehyde removal below regulatory limits (0.1 µg L^{-1}) would take place in a time frame of less than 3-6 h (corresponding to the estimated hydraulic retention time in the sand bed for slow sand filters). Therefore, the initial inoculum was increased from 1 x 10^7 to 4 x 10^7 cells mL^{-1}. The removal of metaldehyde for the selected strains under these conditions is presented in Figure 4-4.

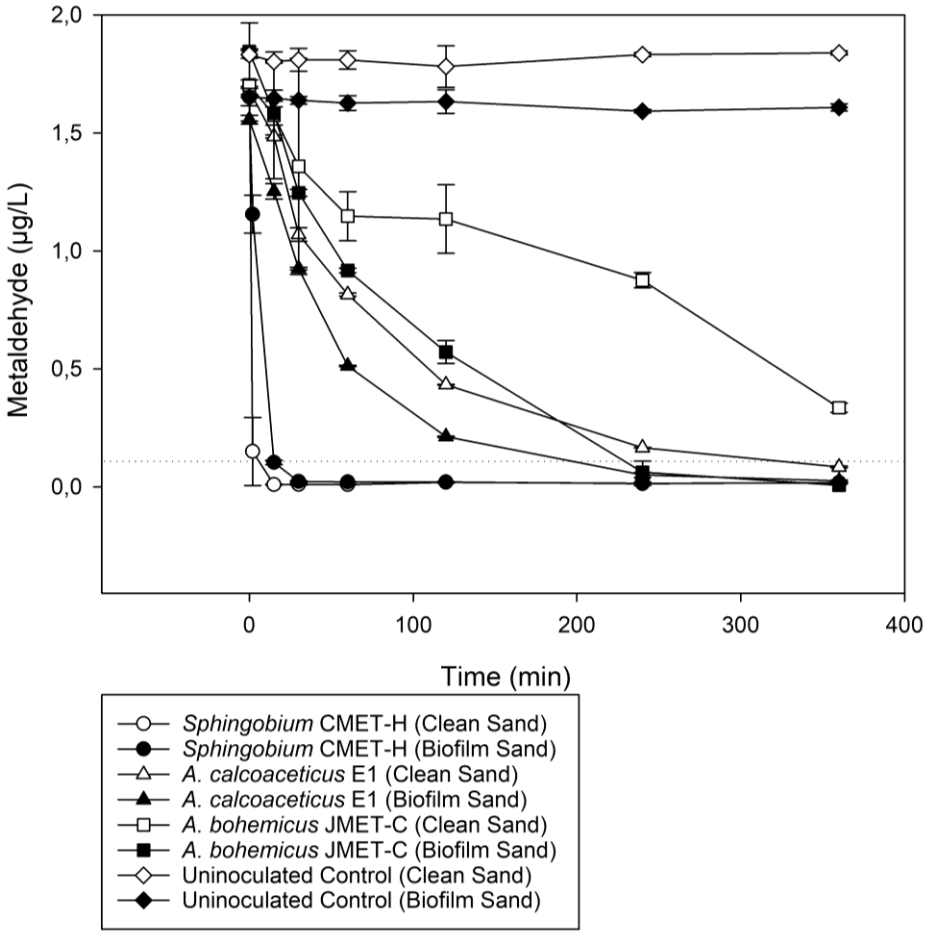


Figure 4-4. Metaldehyde removal assays for bioaugmented raw water and sand systems (4 x 10^7 degrading cells mL^{-1} added) at an environmentally relevant initial concentration (2 µg L^{-1}).The dashed line indicates the regulatory limit (0.1 µg L^{-1}).

Sphingobium CMET-H was able to bring metaldehyde concentrations to levels lower than the regulatory limit in less than 15 minutes for the clean sand systems and in less than 30 minutes for the systems in which sand had a biofilm. *A. calcoaceticus* was able to bring metaldehyde concentrations below the regulatory limit in under 4 h for the sand system with a biofilm and under 6 h for the clean sand system. *A. bohemicus* was able to achieve this in less than 4 hours in the sand system with a biofilm, but not so in the clean sand system. These results show that even in these challenging conditions, with mixed indigenous microbial communities and multiple substrates in a complex matrix it is possible to bring metaldehyde levels below the regulatory limits within a reasonable time frame using a practical inoculum level.

Therefore, two strains were chosen for inoculation of pilot scale slow sand filters: *A. calcoaceticus* E1 and *Sphingobium* CMET-H. Strain E1 was used first because the genes involved in metaldehyde catabolism for this strain had been pinpointed and described (Chapter 2), as well as the mobile genetic elements in charge of its transmission (Chapter 3), whereas for strain CMET-H no verified information regarding the degrading genes was available.

4.3.3 Pilot scale slow sand filter assays

4.3.3.1 Metaldehyde removal

Six pilot scale slow sand filters (22.8 L total volume each) were set up with a continuous flow (1.8 L h^{-1}) of raw water, either spiked (filters 3-6) or not spiked (filters 1 and 2) with metaldehyde at a concentration of 2.0 µg L^{-1}.and a hydraulic retention time of 3.5 h. Metaldehyde removal was quantified through time and is shown in Figure 4-5. A 95% metaldehyde removal would be required to obtain water below the regulatory limit (0.1 µg L^{-1}).

In Phase 1 of the assay (days 1 to 55) metaldehyde removal in the non-bioaugmented (control) slow sand filters (5 and 6) was very low, with no removal detected at most time points and with a maximum removal of 12.3% at time 42.8 d in filter 5 and 11.2% at time 51 d in filter 6.

In filters 3 and 4 *A. calcoaceticus* E1 was first inoculated at time 16 d. In the next 12 h metaldehyde removal peaked at 37.3% in filter 3 and at 56.1% in filter 4. However, the effect did not last for long and the removal percentages returned to pre-inoculation levels at times 18 d for filter 3, and 20 d for filter 4. At time 42 d, double the amount of *A. calcoaceticus* E1 inoculum was added to filters 3 and 4. This addition generated peaks in removal at time 43 d of 74.6% and 78.1% respectively, the highest levels achieved with this strain. Metaldehyde elimination returned to pre-inoculum levels at time 45 d for filter 3 and at time 51 d for filter 4, therefore, the effect lasted for longer than in the first inoculation. Removal levels did not reach the required 95% to obtain compliant water up to this point.

Given these results, in Phase 2 of the assay two different strategies were applied aiming to enhance metaldehyde elimination. Triple the original amount of *A. calcoaceticus* E1 was added to filter 3 only, while inoculation with the strain *Sphingobium* CMET-H was carried out in filters 4 and 5. Filter 4 had been previously treated with *A. calcoaceticus* E1, while filter 5 had not been subjected to previous bioaugmentation. Unexpectedly, after inoculation at time 56 d, filter 3 reached a peak of only 49.1% removal at time 60 d, nonetheless the effect seemed to last for longer than before, not having reached pre-inoculation levels at the end of the assay (16.2% at 72 d).

Remarkably, bioaugmentation with the strain *Sphingobium* CMET-H in filters 4 and 5 at time 56 d led to a fast and drastic increase in the elimination of metaldehyde, achieving complete removal of the compound (below 0.01 μg L^{-1}) at time 58 d for filter 4 and 56.8 d for filter 5. Complete elimination of the compound persisted until time 72 d when the sand filters were shut down and the assay ended. The maximum metaldehyde removal capacity that was observed for this strain occurred shortly after inoculation in sand filter 5 (0.67 μg L^{-1} h^{-1}). Thereafter, since complete elimination was achieved in filters 4 and 5, removal capacity for metaldehyde could be calculated to be 0.57 μg L^{-1} h^{-1}.

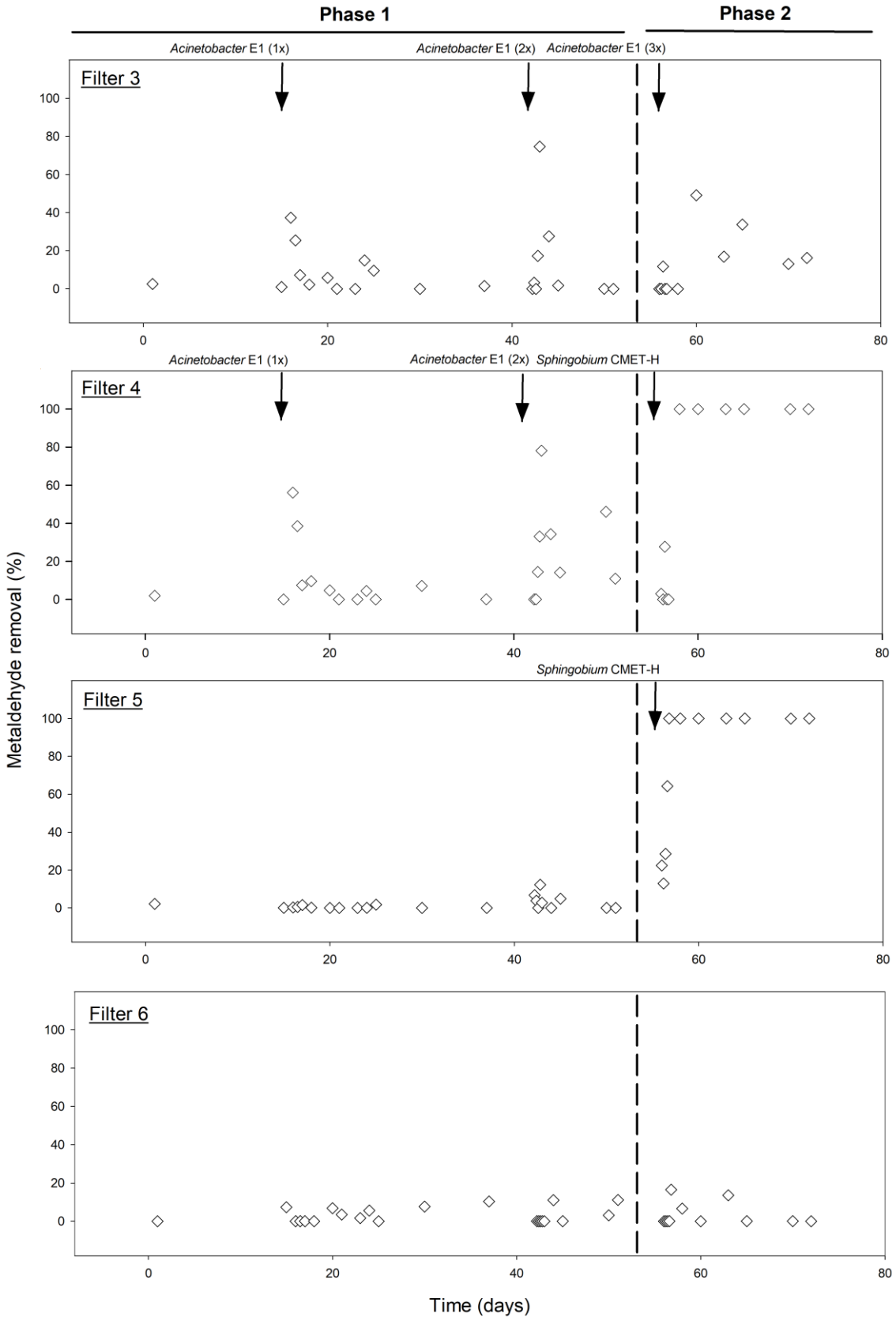


Figure 4-5. Percentages of metaldehyde removal referred to the inlet concentration in slow sand filters with metaldehyde addition (3 to 6). Arrows indicate the moment of inoculation of bioaugmentation agents.

4.3.3.2 Changes in the microbial community

Sequencing quality and depth

Sequencing of 16S rRNA gene amplicons was used to determine the structure of the bacterial and archaeal community of the upper layer of the pilot scale slow sand filters (Schmutzdecke), where most of the biological activity takes place (Ranjan and Prem 2018). Quality statistics for sequencing runs are shown in Annex 3. A total of 78 samples were sequenced, obtaining an average of 303 878 reads per sample and an average of 173 142 reads per sample after quality control and chimera removal.

The sequencing corresponding to the sample from filter 5 at time 23 d was not used for further analyses because a very high number of reads had to be discarded after quality control (only 20 779 reads remained) and very low diversity was identified in the rarefaction plots (Annex 5) when compared to samples from the same filter. Additionally, a significantly different microbial community structure was obtained in this sample when compared to all the other samples, so a technical problem during the sequencing process was assumed to have occurred. All remaining samples were rarefied to a total of 61 312 reads. Rarefaction curves (Annex 5) indicated that for all remaining samples the richness of bacterial taxa had already reached a plateau at this sequencing depth.

Persistence of the bioaugmentation strains

Firstly, an assessment of the relative amounts of the bioaugmentation agents in the upper layer of the slow sand filters was performed. The percentages of the community corresponding to the genera *Acinetobacter* and *Sphingobium* for the filters that underwent bioaugmentation (2 to 5) is shown in Figure 4-6.

A Python script was developed (J. Moir) which allowed the comparison of the 16S rRNA sequence for strains *A. calcoaceticus E1* and *Sphingobium* CMET-H with the reads files for the sand filter sequencing. It was determined that 87.8% of all *Acinetobacter* raw reads had at least

a 99% sequence similarity with *A. calcoaceticus* E1, including 66.9% exact matches. For *Sphingobium* CMET-H, it was determined that 52.6% of all *Sphingobium* raw reads had a 99% sequence similarity with *Sphingobium* CMET-H, including 37.4% of exact matches.

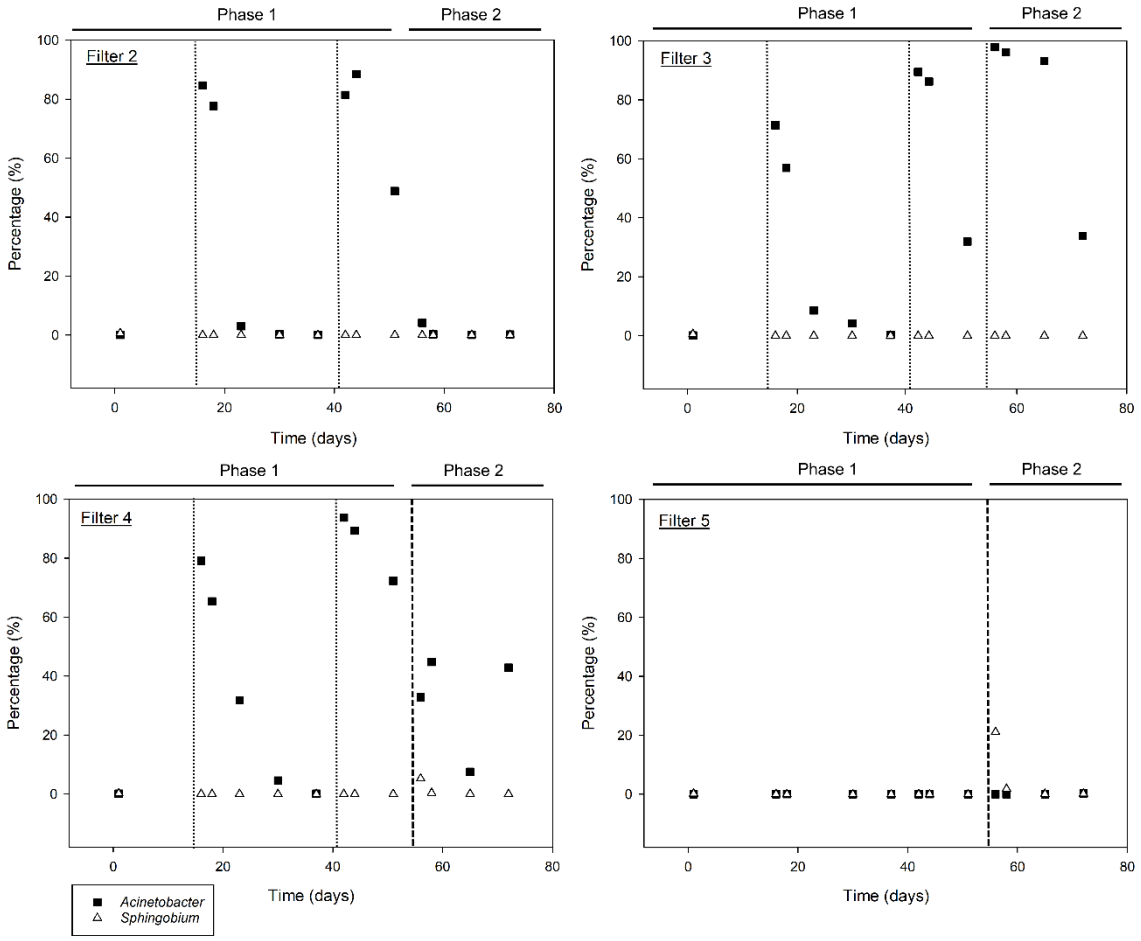


Figure 4-6. Percentage of the microbial population at genus level composed of the genera *Acinetobacter* (filled squares) and *Sphingobium* (empty triangles) in pilot-scale slow sand filters throughout time according to 16S rRNA gene amplicon sequencing. Dotted vertical lines indicate time of inoculation with *A. calcoaceticus* E1 and while dashed vertical lines indicate time of inoculation with *Sphingobium* CMET-H.

For phase 1, before inoculation with *Acinetobacter* E1 in sand filters 2, 3 and 4, the genus was absent at this sequencing depth in filter 2, while the percentage was 0.1% for both filters 3 and 4. After the first bioaugmentation (1x) at time 16 d, the *Acinetobacter* population peaked in the next sampling at 84.6%, 71.4% and 79.0% for filters 2, 3 and 4 respectively. Nonetheless, the

population returned to pre-inoculation levels at times 30-37 d regardless of the presence of metaldehyde in the inlet.

The second bioaugmentation with this same strain (2x) at time 42 d resulted in higher peaks 88.4-93.7% and the percentage of the genus in the community decreased more slowly with time. However, no marked differences in the persistence of the inoculum were observed between the bioaugmented filters that had metaldehyde input (filters 3 and 4) and the one that did not (filter 2.)

Roughly triple the original amount of strain *A. calcoaceticus* E1 was added to filter 3 for phase 2 of the experiment at time 56 d, reaching a peak of 97.9% of the community, nonetheless, at the end of the assay at time 72 d, the population had already dropped to 33.8%.

Overall, these data for *Acinetobacter* indicate that even though a larger inoculum appeared to slightly extend the persistence of the bioaugmentation agent and produce higher peaks in abundance, the elevations on the population were transient and metaldehyde input from the feed water did not appear to be sustaining a sufficiently large and stable degrader population of strain E1.

To better visualise the correlation between the *Acinetobacter* population and metaldehyde removal, these two variables were plotted simultaneously (Figure 4-7). Peaks in metaldehyde removal coincided in time with peaks in the *Acinetobacter* population percentage. However, percentages of at least 71-79% of *Acinetobacter* in the community were necessary to achieve considerable levels (>25%) of metaldehyde removal.

Also on phase 2, *Sphingobium* CMET-H was added to filters 4 and 5 to reach a calculated cell concentration of 5.0×10^7 in the aqueous phase of both filters, a cell number comparable with the 2x treatment with *Acinetobacter* strain E1. Before inoculation, an average population of only 0.02 and 0.03% of members of the genus *Sphingobium* was present in these two sand filters respectively. Peaks of 5.3 and 21.1% of the microbial population in the upper layer of filters 4 and 5 were immediately reached, but decayed rapidly to much lower levels;

nonetheless, as mentioned before, metaldehyde removal was complete and persistent until the end of the assay.

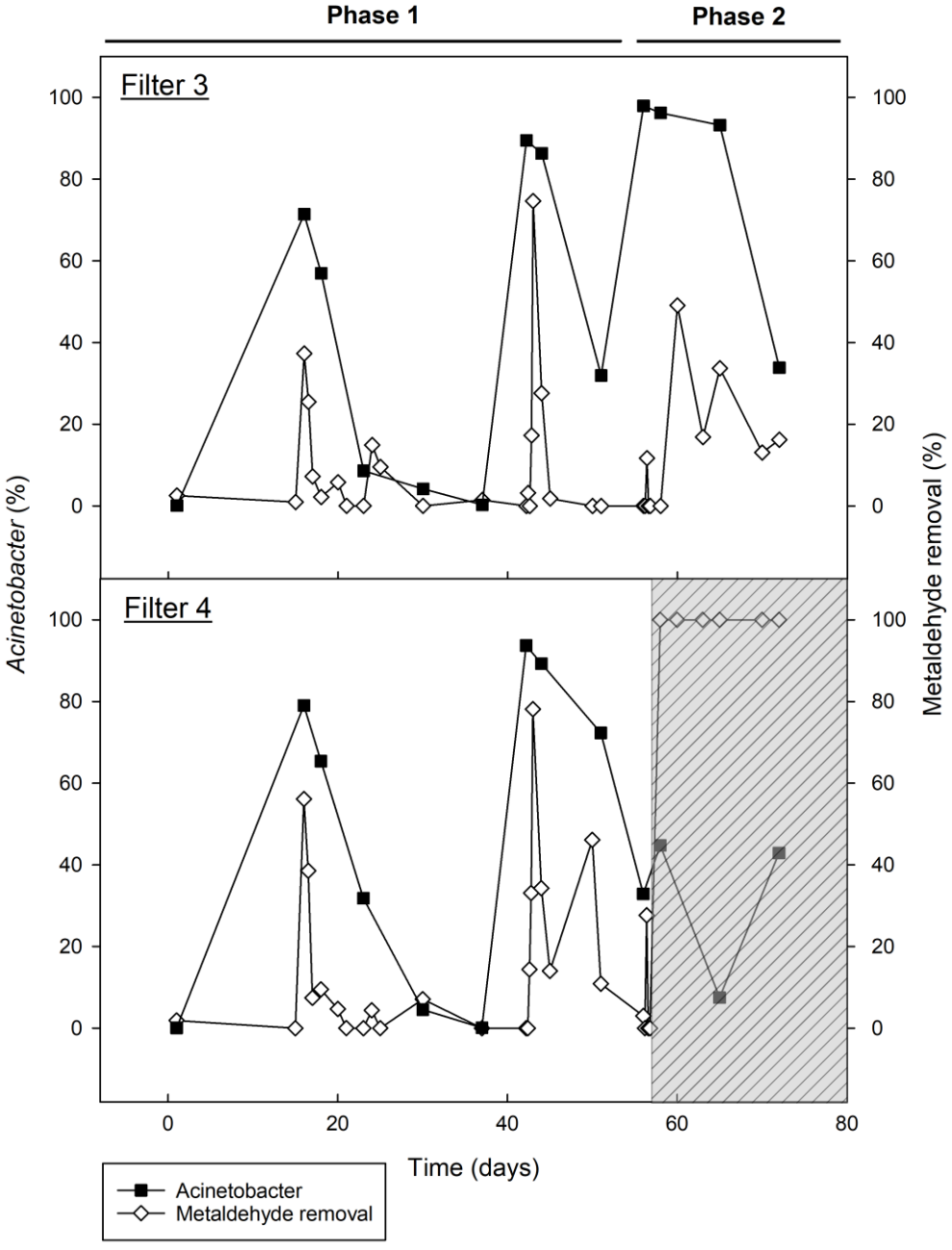


Figure 4-7. Percentage of the bacterial and archaeal community of the slow sand filters composed of genus *Acinetobacter* (filles squares) and metaldehyde removal (empty diamonds) through time. The fraction of the plot corresponding to inoculation with *Sphingobium* CMET-H has been shaded.

Whole microbial community analysis

The rarefied dataset for all filters contained 37 different phyla, 105 classes and 528 genera of Bacteria and Archaea. Figure 4-8a shows a principal coordinates analysis of the whole community at genus level highlighting different filters and time points. PCO1 and PCO2 accounted for 54.8% of the total variation, so a good representation of the community was achieved. Not surprising given the high percentages of *Acinetobacter* found in some samples after bioaugmentation is the fact that inoculation with this strain correlates very well with the sample distribution in the plot, as it is evident in Figure 4-8b. A Pearson correlation coefficient > 0.94 was found between the variable *Acinetobacter* and the direction of the vector in the PCO plots. No evident effect of *Sphingobium* addition can be easily observed in the plots.

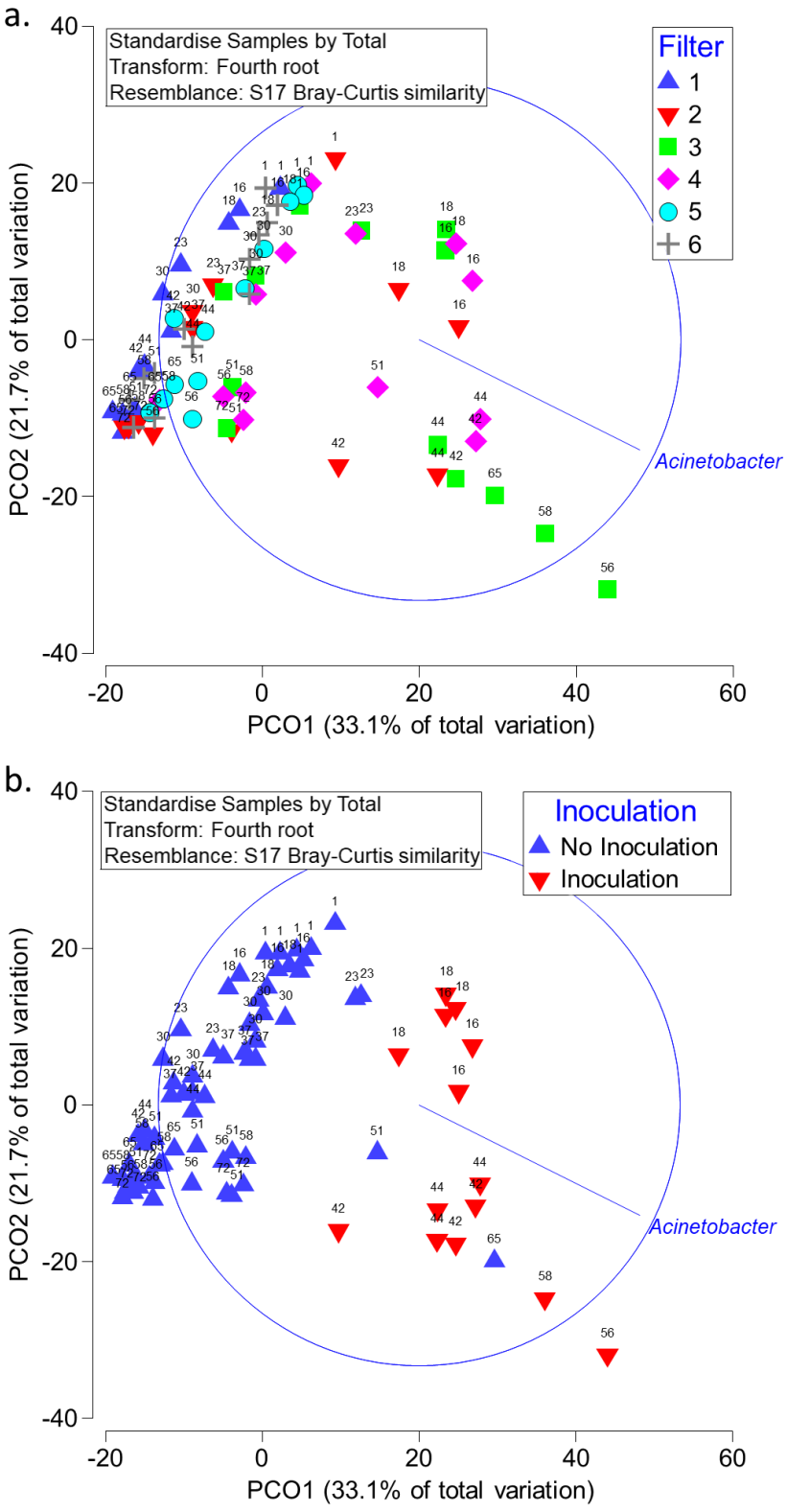


Figure 4-8. Principal coordinates analysis plots at genus level of whole bacterial community 16S rRNA gene amplicons from samples taken from the upper layer of pilot-scale slow sand filters. Labels indicate day of sampling. Samples were rarefied to 61312 sequences each after denoising and quality control. Overlay vectors indicate a Pearson correlation coefficient > 0.94. **a.** Symbols indicate different slow sand filters. **b.** Symbols indicate inoculation (bioaugmentation) treatment with *A. calcoaceticus* E1. Samples were labelled as inoculated on the day of bioaugmentation and on the immediately next time point.

Phylum level analyses

To better visualize the contribution of other taxa to the microbial community, the strains which were artificially added to the community (variables *Acinetobacter* and *Sphingobium*) were removed from the analysis, and the relative abundances of each taxa were recalculated. Phylum level contributions to the microbial community are shown in Figure 4-9. The phylum Proteobacteria tended to heavily dominate the composition of the microbial community in the filters, ranging from 38.3 to 67.7% of the total in the samples. Planctomycetes and Bacteroidetes followed with ranges from 6.1 to 34.5% and from 1.22 to 27.1% respectively.

Data were square root transformed to decrease the influence of the most abundant taxa. PERMANOVA analyses revealed that time of sampling had a significant impact on the community composition at phylum level (Pseudo-F = 5.5488, p = 0.0001, df = 76). Heterogeneity of sample dispersion was discarded by PERMDISP analysis (p = 0.1125).

The biggest average dissimilarity was found between the samples from groups t = 1 d and t = 72 d and was calculated to be 26.22%. A SIMPER analysis revealed that the main contributors to this dissimilarity were phylum Planctomycetes (10.25% relative contribution), bacteria with unassigned phylum (9.32%), phylum Bacteroides (9.03%) and Chloroflexi (8.47%). A shade plot was constructed without Proteobacteria to better visualize these changes in abundance (Figure 4-10). The plot clearly shows that phyla Planctomycetes and Chloroflexi increase their presence with time, while the relative abundance of Bacteroidetes and bacteria with unassigned phylum decreases.

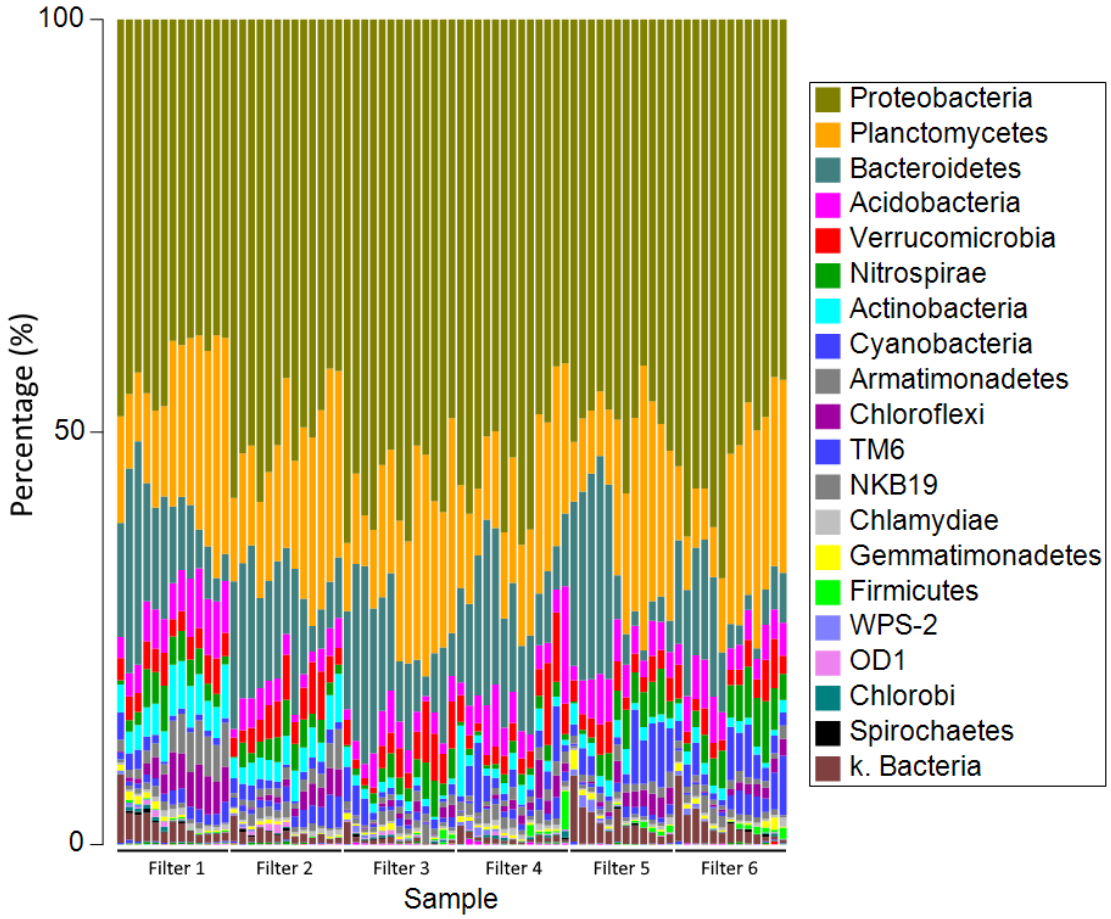


Figure 4-9. Phylum level relative abundance of 16S rRNA gene copies in the top layer of pilot-scale slow sand filters through time. Bioaugmentation agents (*Acinetobacter* and *Sphingobium*) have been removed from the analysis. Different colours indicate different phyla and are ordered in decreasing abundance on the list (restricted to the top 20 taxa). Samples are ordered in chronological order for each filter.

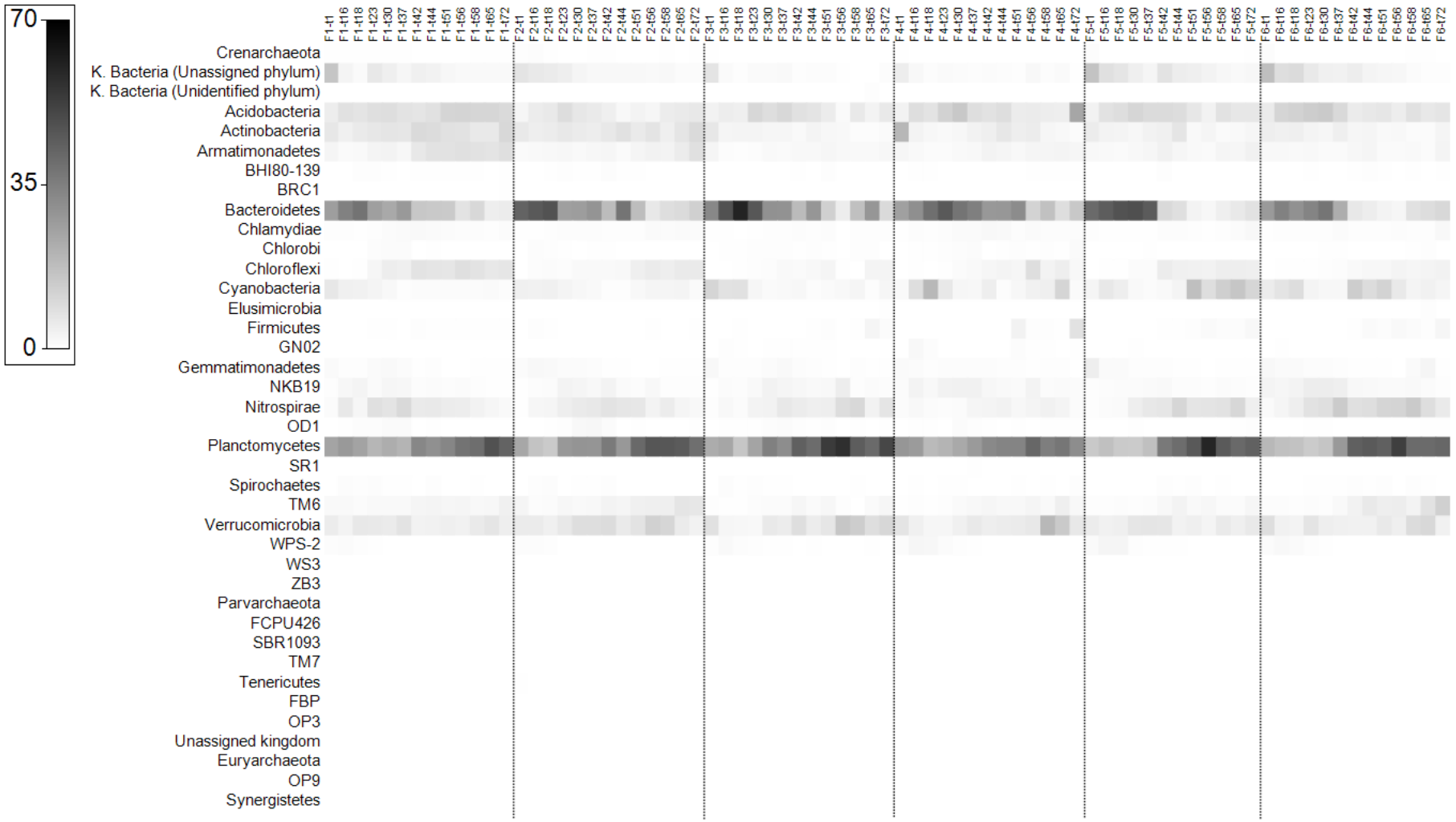


Figure 4-10. Shade plot of phylum level relative abundance for Bacteria and Archaea in slow sand filter samples throughout time. The most abundant phylum, Proteobacteria, was removed from the analysis to better observe the changes in less abundant phyla.

Genus level analyses

As before, the genera *Acinetobacter* and *Sphingobium* were removed from the dataset. The most abundant genera, on average, for all sand filter samples are shown in Table 4-4. Genera inside the phyla Proteobacteria dominated the list, followed by Planctomycetes.

Relative abundance data was transformed by square root. The resulting PCO plots at genus level are shown in Figure 4-11. Here PCO1 and PCO2 accounted for 47.5% of the total variation. Figure 4-11a highlights the new distribution by filter of origin. A PERMANOVA analysis indicated it has an influence on the composition of the community of the top layer of the sand filter (Pseudo-F = 4.6786, p = 0.0001, df = 76). However, it appears to be related with the vertical axis of the plot, which accounts for only 9.3% of the total variation. Stronger correlations with community changes could be observed with the time of sampling (Figure 4-11b). A significant PERMANOVA result was also obtained (Pseudo-F = 5.0677, p = 0.0001, df = 76) and differences between time points in the horizontal axis (38.2% of the total variation) can be clearly observed. Sample group multivariate dispersions were not heterogeneous when analysing these two factors as revealed by PERMDISP results (p = 0.7973 and p = 0.5909).

Table 4-4. Most abundant genera according to relative average abundance across all slow sand filter samples. Genera *Acinetobacter* and *Sphingobium* were removed from the calculations.

Phylum	Class	Order	Family	Genus	Abund. (%)
Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Rhodobacter	4.7
Proteobacteria	Betaproteobacteria	Methylophilales	Methylophilaceae	Methylotenera	4.6
Planctomycetes	Planctomycetia	Planctomycetales	Planctomycetaceae	Planctomyces	4.2
Proteobacteria	Deltaproteobacteria	Myxococcales	Undefined	Undefined	3.9
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Unassigned	3.3
Bacteroidetes	Saprospirae	Saprospirales	Saprospiraceae	Undefined	3.2
Planctomycetes	Planctomycetia	Pirellulales	Pirellulaceae	Undefined	2.6
Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	2.5
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Undefined	Undefined	2.3
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Hydrogenophaga	2.1
Planctomycetes	Planctomycetia	Pirellulales	Pirellulaceae	A17	1.9
Proteobacteria	Alphaproteobacteria	Rhizobiales	Undefined	Undefined	1.9
Unassigned	Unassigned	Unassigned	Unassigned	Unassigned	1.6
Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	Gemmata	1.6
Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Undefined	1.6
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax	1.5
Bacteroidetes	Flavobacteriia	Flavobacteriales	Cryomorphaceae	Undefined	1.4
Proteobacteria	Alphaproteobacteria	Rhodobacterales	Hyphomonadaceae	Woodsholea	1.4
Planctomycetes	Planctomycetia	Gemmatales	Isosphaeraceae	Undefined	1.4
Acidobacteria	Chloracidobacteria	RB41	Ellin6075	Undefined	1.4

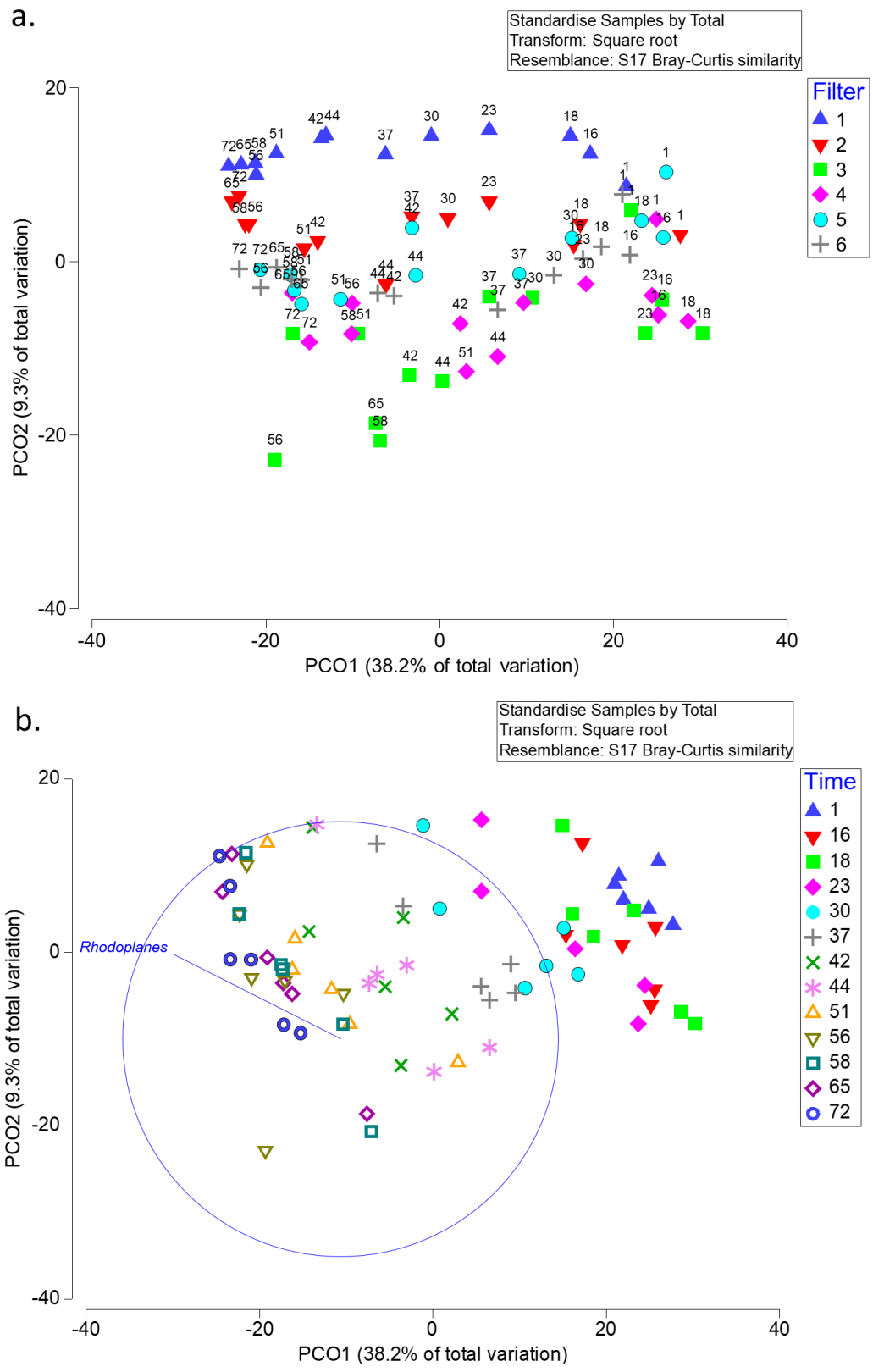


Figure 4-11. Principal coordinates analysis plots at genus level of bacterial community 16S rRNA gene amplicons after removing reads from bioaugmentation agents (*Acinetobacter* and *Sphingobium*). Samples were taken from the upper layer of pilot-scale slow sand filters. **a.** Symbols indicate different slow sand filters, labels indicate day of sampling. **b.** Symbols indicate day of sampling. Samples had been previously rarefied to 61312 sequences each after denoising and quality control.

Genus *Rhodoplanes* was identified as having a Pearson correlation of 0.86 to the vector direction indicated in the plot, which partially coincides with the direction of increasing aging of the samples.

Regarding the different time points, the largest average dissimilarity was once again found between the samples from groups t = 1 and t = 72 d and was calculated to be 54.17%. A SIMPER analysis was performed, and the results for the genera that contribute the most to dissimilarity are shown in Table 4-5. As before, taxa from the phylum Proteobacteria dominated the list. Not surprisingly, several of them coincide with the most abundant genera listed in Table 4-4, nonetheless, others like *Aquabacterium*, an undefined genus of the order Sphingomonadales and an undefined genus of the order Gemmataceae do not. The same trend as in the phylum level analysis is also apparent in the table: genera from the phylum Planctomycetes usually increase in abundance, while genera from the phylum Bacteroidetes decrease. Table 4-6 shows that amongst the best indicators of group differentiation, genera of the family Pirellulaceae, phylum Planctomycetes, are particularly important. Interestingly, genus *Rhodoplanes* appears again, this time as a good indicator of group t = 72 d membership.

The PCO distribution in which the factor addition of metaldehyde is highlighted is shown in Figure 4-12. Even though the PERMANOVA result is significant (Pseudo-F = 6.5425, p = 0.0001, df = 76) with an homogeneous dispersion of the samples as revealed by PERMDISP (p = 0.8729) and there seems to be a separation on the vertical axis for samples exposed and not exposed to metaldehyde, vertical axis variation is only 9.3% of the total. When an ANOSIM analysis is performed with the same data, an R value of 0.206 is obtained (p = 0.0002), which indicates that the sample groups for the metaldehyde addition factor are not separated (Clarke et al. 2014). Given that the factor filter of origin can be having a simultaneous effect on the distribution, the effect of metaldehyde should be interpreted with care, and if any separation in the sample distribution is being caused by the addition of metaldehyde it would be probably be very small.

Table 4-5. Similarity of percentages (SIMPER) analysis showing the bacterial taxa that contribute most to the differences (Av. Diss) in square root transformed relative abundance between the initial and final time point samples for the slow sand filters (Average dissimilarity = 54.17%).

Phylum	Class	Order	Family	Genus	Av. Abun t=1 (%)	Av. Abun t=72 (%)	Av. Diss	Diss/ SD	Contrib. (%)	Cumul. (%)
Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Rhodobacter	2.8	1.1	0.8	4.3	1.5	1.5
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Aquabacterium	1.7	0.1	0.8	1.0	1.5	3.0
Bacteroidetes	Saprospirae	Saprospirales	Saprospiraceae	Undefined	2.2	0.5	0.8	3.5	1.5	4.5
Planctomycetes	Planctomycetia	Pirellulales	Pirellulaceae	A17	0.2	1.8	0.8	9.4	1.4	5.9
Unassigned	Unassigned	Unassigned	Unassigned	Unassigned	2.2	0.8	0.7	2.3	1.3	7.2
Planctomycetes	Planctomycetia	Gemmatales	Isosphaeraceae	Undefined	0.1	1.6	0.7	2.9	1.3	8.5
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Unassigned	2.4	0.9	0.7	3.0	1.3	9.8
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax	1.7	0.4	0.7	2.0	1.2	11.0
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Undefined	Undefined	0.9	2.3	0.6	1.5	1.2	12.2
Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	Undefined	0.3	1.6	0.6	2.3	1.2	13.3

Table 4-6. Similarity of percentages (SIMPER) analysis showing the bacterial taxa that are the best indicators of group differences (Diss/SD) in square root transformed relative abundance between the initial and final time point samples for the slow sand filters (Average dissimilarity = 54.17%).

Phylum	Class	Order	Family	Genus	Av. Abun t=1 (%)	Av. Abun t=72 (%)	Av. Diss	Diss /SD	Contrib. (%)
Planctomycetes	Planctomycetia	Pirellulales	Pirellulaceae	A17	0.2	1.8	0.8	9.4	1.4
Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	Unassigned	0.1	0.8	0.3	6.5	0.6
Bacteroidetes	Cytophagia	Cytophagales	Cytophagaceae	Leadbetterella	0.4	0.0	0.2	6.1	0.4
Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	Salinibacterium	0.7	0.2	0.2	5.9	0.4
Planctomycetes	Planctomycetia	Pirellulales	Pirellulaceae	Unassigned	0.9	2.0	0.5	5.3	1.0
Planctomycetes	OM190	CL500-15	Unassigned	Unassigned	0.9	0.1	0.4	5.2	0.7
Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Rhodoplanes	0.3	0.7	0.2	4.9	0.3
Chloroflexi	TK10	AKYG885	Dolo_23	Unassigned	0.1	0.8	0.3	4.5	0.6
WPS-2	Unassigned	Unassigned	Unassigned	Unassigned	0.6	0.1	0.3	4.5	0.5
Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Phenylobacterium	0.9	0.2	0.3	4.4	0.6

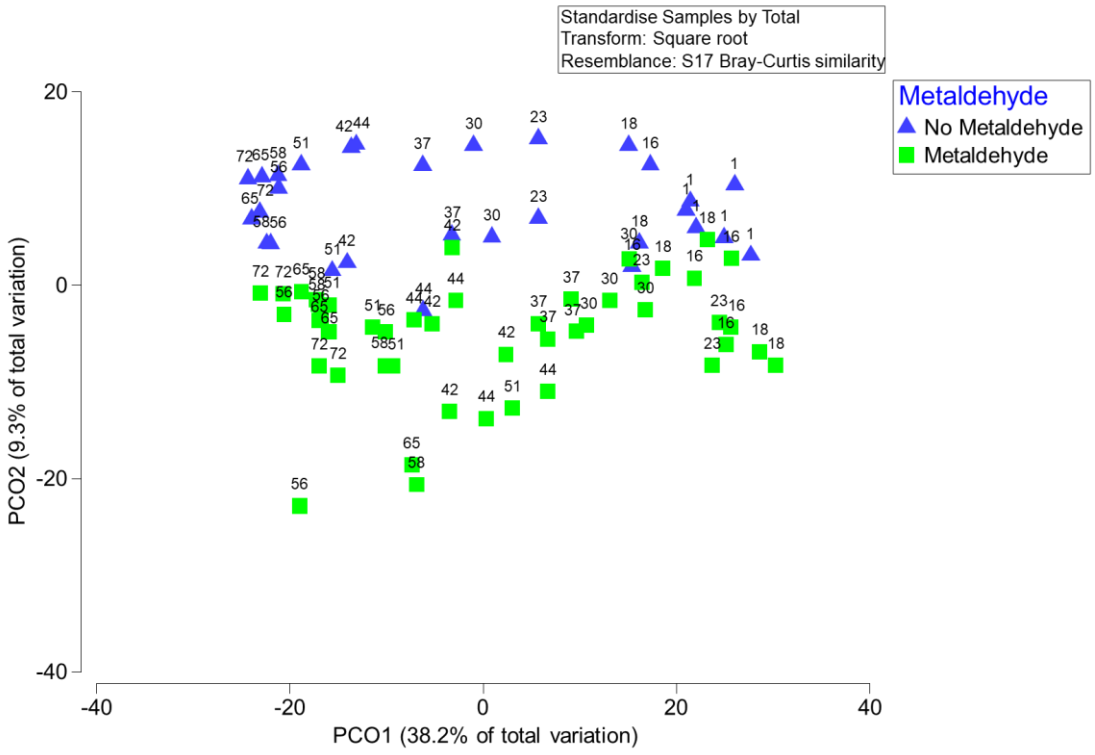


Figure 4-12. Principal coordinates analysis plot at genus level of bacterial community 16S rRNA gene amplicons after removing reads from bioaugmentation agents (*Acinetobacter* and *Sphingobium*). Samples were taken from the upper layer of pilot-scale slow sand filters. Symbols indicate presence or absence of metaldehyde (2 µg L^{-1}) in the inlet water; labels indicate day of sampling. Samples had been previously rarefied to 61312 sequences each after denoising and quality control.

Since all previous metaldehyde-degrading isolates have been classified as Proteobacteria, all other taxa were removed from the data in an attempt to obtain additional information regarding significant changes at genus level related to metaldehyde addition to the inlet water. Nonetheless, the PERMANOVA (Pseudo-F = 8.3570, p = 0.0001, df = 76) and ANOSIM (R = 0.256) statistics, as well as the PCO ordination (Annex 6), did not change significantly.

Microbial diversity through time in each individual sand filter at genus level was measured using the Shannon Diversity Index (H') and the results were plotted in Figure 4-13. A trend of increasing diversity with time can be recognized in the microbial community of the sand filters, however no marked differences were present between individual sand filters.

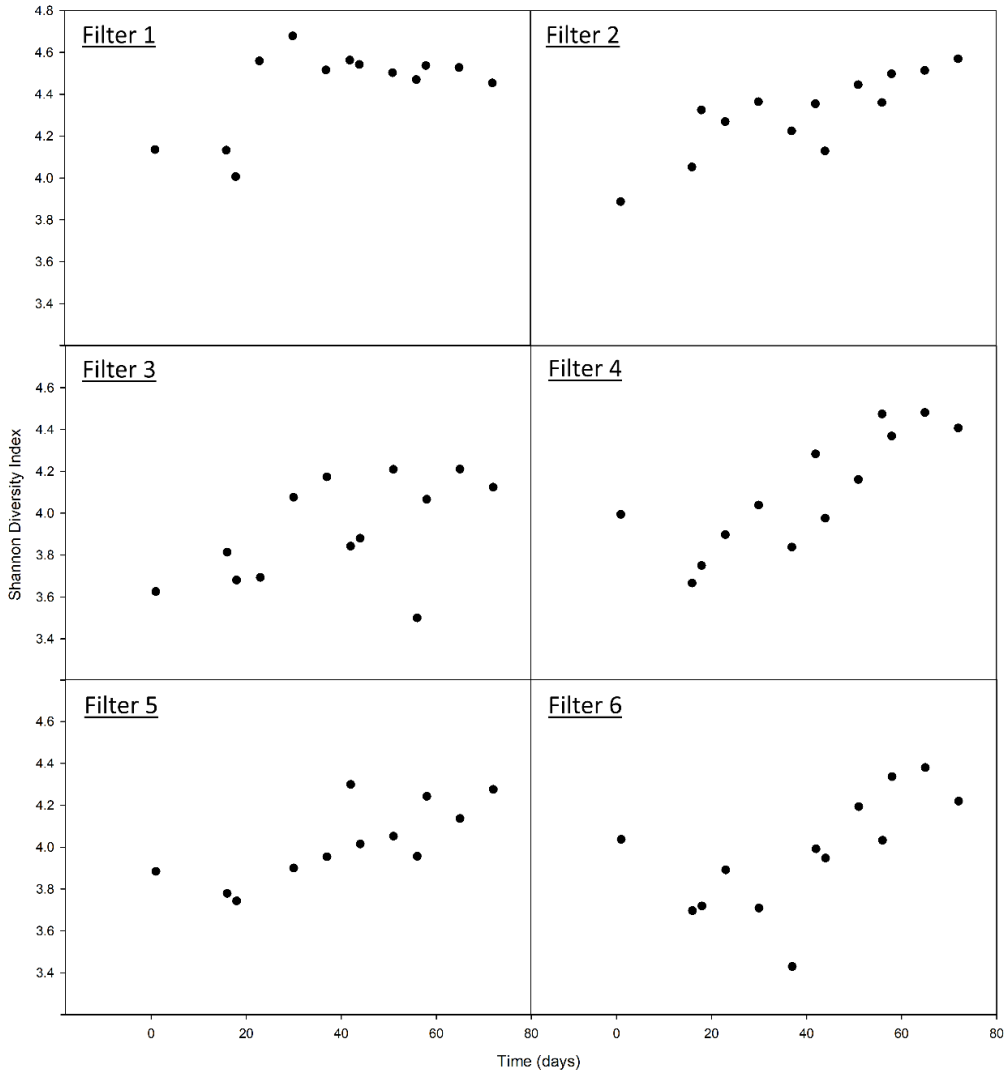


Figure 4-13. Microbial diversity measured by Shannon Diversity Index (H') at genus level in slow sand filter samples throughout time. Reads from bioaugmentation agents (*Acinetobacter* and *Sphingobium*) have been removed from the analysis.

4.4 Discussion

4.4.1 Identification of the most suitable bioaugmentation candidates

Autochthonous microbial communities in biologically-active sand filters seldom contain dominating populations capable of mineralizing specific micropollutants that can be present in drinking water sources at the ng L^{-1} to µg L^{-1} range (Benner et al. 2013). This is the rationale behind the idea of using bioaugmentation with specialist degraders for water purification.

To be able to ultimately apply bioaugmentation to slow sand filters, the first objective of this chapter was to identify the most suitable bioaugmentation candidates from a library of metaldehyde-degrading strains through bench-scale assays. In general, most potential bioaugmentation agents have only been shown to be able to mineralize and grow on the target contaminants at much higher concentrations than the ones associated with micropollutants (Benner et al. 2013).Therefore, it was of particular importance to confirm substrate utilization at very low concentrations especially given the fact that the degrading strains had been isolated by enrichment with initial concentrations of metaldehyde up to 50 000 times higher (100 mg L^{-1}) than the starting concentrations in this assay (2.0 µg L^{-1}).

First, the removal of metaldehyde by the degrading strains in pure culture and defined medium to levels below the regulatory limit was tested and it was found to be successful for three of the strains within the desired time frame. A positive feature is that these strains have been shown to use the contaminant as a growth substrate (Chapter 2) in metabolic reactions and therefore are able to mineralize it to innocuous compounds, so no persistent and potentially toxic transformation products are likely to be generated.

Other strains (*A. lwoffii* SMET-C and *P. vancouverensis* SMET-B), however, did not show an ability to remove the contaminant at this low starting concentration, even though they share the metaldehyde-degrading gene cluster with two of the successfully degrading strains *A. calcoaceticus* E1 and *A. bohemicus* JMET-C.

Inability to degrade contaminants at lower concentrations can be been linked to two main drivers: either a limited bioavailability of the contaminant that would prevent an otherwise swift enzymatic transformation (Bosma et al. 1996), or the downregulation of catabolic enzymes at low concentrations of substrates (Kundu et al. 2019).

In the first scenario a combination of physico-chemical factors such as phase distribution and mass transfer come into play, as well as physiological factors of the microorganisms such as membrane permeability, active uptake systems, enzymatic machinery and excretion of

surfactants and enzymes (Cirja et al. 2008). For the second case it is important to consider that bacteria can prioritize the utilization of a single energy favourable substrate at high concentrations (> 10 mg L^{-1}) by repressing the expression of other catabolic pathways (Kalisky et al. 2007), however, degradation of the contaminant is not prioritized when it is present at µg L^{-1} concentrations (Horemans et al. 2013).

Since the identified pathway of metaldehyde degradation appears to have evolved recently (Chapter 3), it is unlikely that efficient gene expression regulation mechanisms have had a chance to evolve, so its expression would be predominantly constitutive, as has been observed for other pesticides (Bers et al. 2011; Sørensen et al. 2009). For our model metaldehyde degrader, *A. calcoaceticus* E1, this is supported by the observation that the onset of metaldehyde degradation proceeds very rapidly after inoculum addition to a metaldehyde-containing aqueous matrix, even if the inoculum has been pre-grown in nutrient rich media without the contaminant. This has been observed at elevated (100 mg L^{-1}; E. Fuller, personal communication) or low (2.0 µg L^{-1}; this study) concentrations of the pesticide.. Thomas et al. (2017) found a twofold increase in metaldehyde-degrading activity after culturing with metaldehyde as opposed to acetate as a sole carbon source. Overall, these observations indicate that in *A. calcoaceticus* E1, while the pathway is probably constitutive, it can be partially upregulated. If this is also the case for the other strains, then their limited degradation at low concentrations would be likely due to low bioavailability instead of decreased enzyme expression.

The best performing strains were tested at increased inoculum concentrations in more challenging conditions with mixed microbial populations and dissolved organic compounds. The additional carbon sources can have a positive or negative effect on contaminant degradation. Positive effects can be the result of increased microbial biomass; however, this is usually the case in longer experiments that allow for more extensive microbial growth. The negative effects

may occur if specific compounds other than the pesticide are present in high enough concentrations that cause catabolome repression (Kalisky et al. 2007).

All in all, in this part of the study, the degrading strains were capable of bringing metaldehyde levels below the regulatory limit in the required time frame (< 6 h) in this complex matrix that more closely resembles a slow sand filter environment.

4.4.2 Removal of metaldehyde in bioaugmented pilot-scale slow sand filters

The second objective in this chapter was to evaluate if bioaugmentation of upscaled slow sand filter systems with the selected strains is an effective strategy for removing the contaminant. Additions of *Acinetobacter* strain E1 in calculated final concentrations ranging from 2.40 x 10^7 (1x) to 8.11 x 10^7 (~3x) could only reach a maximum average elimination of 76.4%, which occurred in the 2x inoculation, falling short of the 95% removal required to achieve levels below the regulatory limit. Furthermore, removal of the compound was only transient and it quickly returned to pre-inoculation levels.

On the other hand, a single addition of *Sphingobium* CMET-H led to persistently undetectable levels of metaldehyde for approximately two weeks until slow sand filter shutdown. The maximum removal capacity for metaldehyde for this strain was observed after inoculation in sand filter 5 at 0.67 µg L^{-1} h^{-1} from a continuous flow inlet concentration of 2.0 µg L^{-1} and a hydraulic retention time of 3.5 h. Thereafter, removal capacity for metaldehyde could be calculated to be 0.57 µg L^{-1} h^{-1}, since complete elimination persisted.

Comparatively, (Rolph et al. 2019) used smaller laboratory-scale fluidised-bed through-flow columns with previously acclimated sand (50 µg L^{-1} metaldehyde, 5 d pre-exposure) and obtained a maximum removal rate of 0.17 µg L^{-1} h^{-1} from an inlet concentration of 0.5 µg L^{-1} and 13.8 h contact time. Metaldehyde removal was increased by 40% in acclimated vs. non-acclimated columns. They managed to obtain water within compliance limits for more than 20 days. Rolph (2016) used a fluidised bed reactor (100 L volume) with active sand and non-

controlled temperature and reached a maximum removal rate of approximately 0.75 µg L^{-1} h^{-1} initially, however decreasing thereafter to much lower levels from an inlet concentration of 0.1-0.9 µg L^{-1}, with 67 min hydraulic retention time and a recycle ratio of 28.6. Thus, the use of *Sphingobium* CMET-H in this study led to an efficient and stable removal of the compound and achieved compliant water for 14 d with a less energy intensive technology than other biological strategies (fluidized bed reactors). More challenging conditions such as lower temperatures and varying metaldehyde concentrations in the inlet could be tested in the future with this strain.

4.4.3 Fate of metaldehyde degraders in the microbial community of slow sand filters

The third objective was to determine if the presence of an active population of metaldehyde degraders is the key factor governing its removal. The success of a bioaugmentation strategy depends on the ability of the added microorganism to integrate into the microbial community and carry out a metabolic utilization of the contaminant (Benner et al. 2013).

Upon increasing levels of *A. calcoaceticus* E1 inoculum, higher although transient percentages of sand filter microbial community corresponded to the *Acinetobacter* taxon. Only when a remarkably high *Acinetobacter* population was reached did metaldehyde removal was significant, although for this strain it never reached the required 95% elimination to obtain compliant levels in the effluent.

Little or no effect on bioaugmentation strategies may be due to several different reasons, for instance, insufficient microbial population in the inoculum, failure on the adaptation process of the inoculated microorganisms to the specific environmental conditions (oxygen supply, temperature, moisture content, pH, toxicity, etc.), insufficiency of the substrate, low nutrient concentration, competition between the native microorganisms and the introduced inoculum and grazing by protozoa are often cited (Bouchez et al. 2000; R. M. Goldstein et al. 1985; Vidali 2001). Additionally, in sand filters, adhesion of the bioaugmentation agent to the

sand particles is of particular importance to prevent inoculum wash-out and leaching and allow its integration into the biofilm (Samuelsen et al. 2017). Ecological factors such as a high level of microbial community evenness have also been negatively correlated with the ability of invasion by an external microorganism (De Roy et al. 2013).

For *A. calcoaceticus* E1, a combination of factors may have hindered its performance in the sand filters. Even though the strain heavily dominated the community composition at several time points, even reaching levels close to 98% after the 3x inoculation, the required metaldehyde elimination level was not reached, not even transiently. This indicates that in order to achieve this, an even higher inoculum would have been necessary, at least in the order of 10^8 C.F.U. mL^{-1} of water inside the reactor, a magnitude no longer considered practical. Hence, the strain/enzymatic machinery is probably just not efficient enough in the tested slow sand filter conditions. Furthermore, as *Acinetobacter* populations and metaldehyde elimination were only transient, a rapid washout of the inoculum from the sand filters would still be likely.

On the other hand, even though a relatively low proportion of the top layer of the sand filter community corresponded to the taxon *Sphingobium* after *Sphingobium* CMET-H was inoculated (t = 56 d) and the percentage decreased rapidly, metaldehyde removal was complete and persistent in sand filters 4 and 5. This observation can have several different explanations that may be taking place alone or in combination: (1) a very low amount of the strain/enzyme is needed to achieve complete removal of the compound; (2) *Sphingobium* CMET-H is being retained further down in the columns and not in the top layer (Schmutzdecke); and (3) horizontal transfer of the metaldehyde-degrading genes is taking place from *Sphingobium* CMET-H to the indigenous microbiota in the sand filters. Figure 4-14 summarizes these hypotheses and the experimental designs that could be used to put them to test.

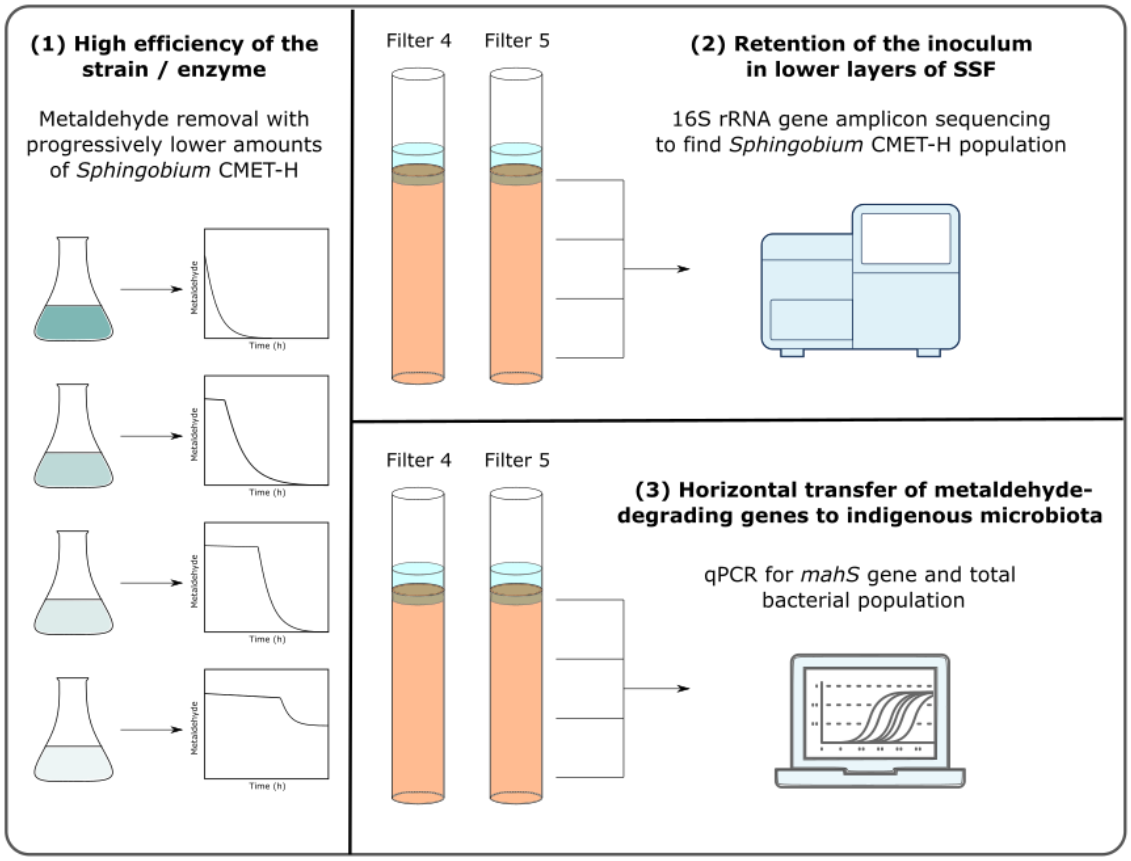


Figure 4-14. Hypotheses for complete metaldehyde elimination in the sand filters despite low abundance of the bioaugmentation strain (*Sphingobium* CMET-H) in the Schmutzdecke and possible experimental designs to test each of them.

To test the first hypothesis (very high enzyme/inoculum efficiency), a laboratory-scale batch assay using Erlenmeyer flasks or a continuous flow assay using small columns (econo-columns) with raw water and sand with a biofilm could be performed. Here, progressive decreasing concentrations of *Sphingobium* CMET-H inoculum would be used. That way, the critical bacterial biomass necessary for metaldehyde elimination could be calculated and extrapolated to the pilot-scale sand filters. A similar experimental design was used by Bourne et al. (2006) to assess the degradation of the cyanobacterial toxin microcystin LR in natural water.

A relatively simple procedure could be followed to test the second hypothesis (inoculum retention in lower levels): samples of sand from the lower layers of the sand filters used in this chapter were collected at time 72 d and frozen (kept at Cranfield University) and could undergo

DNA extraction, 16S rRNA amplification and amplicon sequencing to obtain the percentage of the community composed of *Sphingobium* in the lower layers.

For the third hypothesis (horizontal transfer of degrading capacity), it is important to mention that recently, by comparing the *mahX* degrading gene sequence from *A. calcoaceticus* E1 with the whole genome of *Sphingobium* CMET-H, a candidate gene for the first-step metaldehyde-degrading protein in *Sphingobium* was identified, and heterologous expression in *E. coli* confirmed its role as a metaldehyde-degrading enzyme. It has been provisionally named *mahS* and bioinformatic analysis has predicted it to be harboured in a conjugative plasmid in *Sphingobium* CMET-H (E. Fuller, personal communication). Hence, one way to test for horizontal gene transfer would involve using qPCR to compare the copy numbers of *mahS* genes in the bioaugmented sand samples with the *Sphingobium* population numbers and see if significant differences arise through time. This would suggest horizontal gene transfer of *mahS* to the indigenous microbial population. Oxford Nanopore long-read sequencing could also be used for *Sphingobium* CMET-H, as has been for *A. calcoaceticus* E1 (Chapter 3), to confirm the presence of the conjugative plasmid and the genetic context of the degrading genes.

4.4.4 Whole microbial community analysis

The most abundant phyla in the sand filter samples from this study were, in order, Proteobacteria, Planctomycetes and Bacteroidetes. Other studies have also found Proteobacteria to be the most abundant phylum in slow sand filters, usually followed by Bacteroidetes (D'Alessio et al. 2015; Haig et al. 2014). Hence, there is a higher proportion of Planctomycetes in the Schmutzdecke samples from this study. The total number of bacterial genera found in samples from all six slow sand filters in this study was 528. This is consistent with other studies, which found 688 genera were in the media and water of two full scale slow sand filters (Haig et al. 2015) and 598 genera in laboratory and full scale slow sand filters (Haig et al. 2014).

It was found that filter aging was the most important factor correlating with changes in the microbial community, an observation that has been found in other studies as well (D'Alessio et al. 2015; Haig et al. 2015). The proportion of phylum Bacteroidetes tended to decrease with time. This has been seen in other studies with sand filtration (Mukherjee et al. 2016; Pfannes et al. 2015), and could be explained by preferential grazing by protozoans (Jezbera et al. 2006) or inadequate (aerobic) redox conditions in the upper part of the sand filters (Webster and Fierer 2019). Contrastingly, Planctomycetes increased with aging, a contrasting observation with other studies (Haig et al. 2015) in which they were more abundant in early-stage slow sand filters. SIMPER analysis also indicated that several Planctomycetes were indicators of differentiation between groups t = 1 and t = 72. An increased exposure to light can increase the algal population, which promotes growth of Planctomycetes (Pizzetti et al. 2011). In this study, the water transitioned from being stored in complete darkness to being partially exposed to artificial light. On the other hand, in the abovementioned study there was more exposure to light in the juvenile sand filters, which might explain the observed differences. This may also explain the increase in abundance of the genus Rhodoplanes, a phototrophic Alphaproteobacterium (Hiraishi and Ueda 1994), with filter aging.

Metaldehyde presence in the influent water did not generate evident changes in the microbial community. In this study metaldehyde was used at micropollutant levels. Growth of pure cultures on single substrates requires a certain minimum level of the compound (threshold concentration), which has typically been reported to range between 1-100 μg L^{-1} for different compounds (Egli 2010). Even though this threshold has not been exactly determined for metaldehyde, it is very unlikely that 2.0 μg L^{-1} concentrations of the compound would generate large enough changes in the overall microbial community to be detected confidently. Furthermore, it was not a specific objective of the study to detect these differences, if present. Nonetheless, a non-parametric comparison performed by Rolph et al. (2019) found that the microbial community of slow sand filter samples acclimated with a higher concentration of

metaldehyde (50 µg L^{-1}) was statistically different from the community of non-acclimated samples, even though more detailed microbiological work was suggested to adequately elucidate these differences.

The diversity of the microbial community showed an overall increase through time in all sand filters. Similar patterns have been found in previous studies with slow sand filters, in which the number of OTUs had a strong positive correlation with filter age (Haig et al. 2014; Ramond et al. 2013). Webster and Fierer (2019) also found and increase in richness and evenness of sand filter samples through time. These observations go in hand with an enhancement of sand filter performance with maturity (Kem 1996; Pompei et al. 2017).

4.5 Conclusions

In this chapter, a library of metaldehyde-degrading bacteria was screened using bench-scale assays and the best strains for removal of the compound, at very low starting concentrations and in challenging conditions, were successfully determined to be *A. calcoaceticus* E1 and *Sphingobium* CMET-H.

These strains were used as inoculum for bioaugmentation of pilot-scale slow sand filters. Removal of the compound by increasing amounts of *A. calcoaceticus* E1 was insufficient to achieve compliant water and was restricted to the periods of time in which its population was remarkably high. Low degradation efficiency in the experimental conditions and inoculum washout were the most likely causes of this behaviour. Contrastingly, bioaugmentation with *Sphingobium* CMET-H could swiftly bring metaldehyde concentrations in water to compliant levels that persisted for 14 d until reactor shutdown. Interestingly, even though complete removal was taking place, its population in the top layer of the sand filters was relatively low. Several hypotheses, including a very high efficiency of the enzyme/strain, retention of the strain in lower levels of the sand filter, and horizontal transfer of degrading genes to the indigenous microbiota were proposed as possible explanations.

Once these hypotheses are validated experimentally, valuable information regarding the behaviour of *Sphingobium* CMET-H and its degrading capabilities will be obtained, and together with the important advances in the understanding of the degrading pathway that are currently being made for the strain, its use at an even larger scale for water purification can be devised.

CHAPTER 5: FINAL DISCUSSION

Base knowledge on the biological degradation of metaldehyde was very limited for several decades. It was known that metaldehyde was metabolised by microorganisms producing acetaldehyde in first instance and CO_2 as a final mineralization product (Bieri 2003; Simms et al. 2006), and studies regarding its persistence in soil were undertaken (Cranor 1990; Möllerfeld et al. 1993; X. Zhang and Dai 2006).

It was not until 2007 that adequate method for metaldehyde quantification at very low concentrations was developed; since then, the compound has been found in numerous occasions in groundwater, surface water and processed water (Stuart et al., 2011). Even though the detected levels do not represent an immediate threat to human health, they do generate drinking water quality compliance breaches. Thus, the interest in finding solutions to the metaldehyde problem quickly increased within water companies and in academia. Several potential solutions have been proposed having met with varying degrees of success and cost-effectiveness, including catchment management initiatives, product substitution, a ban on the product, which was legally challenged afterwards, advanced oxidation methods and specially designed materials.

Biological solutions began to be explored and important steps such as the first isolation of metaldehyde-degrading bacterial strains (Thomas et al. 2017) and the study and use of acclimation of microbial communities for metaldehyde removal in slow sand filter systems (Rolph et al. 2019) were taken. More recently, metaldehyde-degrading potential was found to be ubiquitous in a wide range of soils (Balashova et al. 2020). Regardless, knowledge related to the molecular underpinnings of the microbial degradation of metaldehyde was very scarce prior to this thesis. Firmly established knowledge was limited to the work of Thomas (2016) who demonstrated that there was a metaldehyde-dependent upregulation of an acetaldehyde dehydrogenase enzyme in *A. calcoaceticus* E1.

The cornerstone of the work presented here was the initial isolation of numerous and diverse metaldehyde degrading strains. This was the key to several important results: (1) the identification of the metaldehyde-degrading gene cluster through comparative genomics with several strains from different taxa; (2) the elucidation of the mobile genetic elements associated with the horizontal transfer of the gene cluster by examination of their genetic context in different strains; and (3) the successful bioaugmentation of pilot-scale slow sand filters by testing and selecting within a diverse library of strains.

This thesis highlights the continuing value of "traditional" experimental culture-based methods, and their application in coordination with contemporary molecular techniques. After all, the most informative genetic and biochemical analyses can still only be performed on culturable organisms. Therefore, efforts should be made to effectively isolate and culture these environmentally-relevant microbes. This is what allows a great amount of valuable information to be added to genomic and proteomic bioinformatic databases which are, ultimately, the ones used to assign gene functions in studies that employ newer techniques such as whole-community metagenomics.

Notwithstanding the value of traditional methods, much more sophisticated techniques were used, such as third generation sequencing, which was vital to the identification of plasmids in the model degrader *A. calcoaceticus* E1, and quantification of metaldehyde by LC-MS/MS, which provided the level of sensitivity required to adequately test the strains against environmentally-relevant concentrations of the contaminant.

Well established methods like qPCR were used to confirm that the metaldehyde-degrading population proliferates in response to the pesticide and that higher number of gene copies correlated with periods of metaldehyde removal in several soils, while next-generation sequencing provided valuable data on the fate of the microorganisms used for bioaugmentation and on the overall microbial community structure.

Even so, several lines of work are still ahead, including: the elucidation of alternative metaldehyde degradation pathways that exist in nature; the development of biosensors for metaldehyde; the study of the impact of different experimental conditions in bioaugmented slow sand filters, with periods of metaldehyde starvation and lower operation temperatures; and the scaling up and engineering larger bioremediation systems for metaldehyde.

References

Abraham, J., & Silambarasan, S. (2013). Biodegradation of chlorpyrifos and its hydrolyzing metabolite 3,5,6-trichloro-2-pyridinol by Sphingobacterium sp. JAS3. *Process Biochemistry*, *48*(10), 1559–1564. https://doi.org/10.1016/J.PROCBIO.2013.06.034

Addou, S., Rentzsch, R., Lee, D., & Orengo, C. A. (2009). Domain-Based and Family-Specific Sequence Identity Thresholds Increase the Levels of Reliable Protein Function Transfer. *Journal of Molecular Biology*, *387*(2), 416–430. https://doi.org/10.1016/J.JMB.2008.12.045

Albers, C. N., Feld, L., Ellegaard-Jensen, L., & Aamand, J. (2015). Degradation of trace concentrations of the persistent groundwater pollutant 2,6-dichlorobenzamide (BAM) in bioaugmented rapid sand filters. *Water Research*, *83*, 61–70. https://doi.org/10.1016/j.watres.2015.06.023

Aldén, L., Demoling, F., & Bååth, E. (2001). Rapid method of determining factors limiting bacterial growth in soil. *Applied and Environmental Microbiology*, *67*(4), 1830–8. https://doi.org/10.1128/AEM.67.4.1830-1838.2001

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Anderson, M. J., Gorley, R. N., & Clarke, K. R. (2008). *PERMANOVA+ for PRIMER: Guide to Software and Statistical Methods*. Plymouth, UK: PRIMER-E.

APHA-AWWA-WEF. (2012). *Standard Methods for the Examination of Water and Wastewater* (22nd ed.). Washington.

Arbeli, Z., & Fuentes, C. L. (2007). Accelerated biodegradation of pesticides: An overview of the phenomenon, its basis and possible solutions; and a discussion on the tropical dimension. *Crop Protection*, *26*(12), 1733–1746. https://doi.org/10.1016/J.CROPRO.2007.03.009

Arredondo-Alonso, S., Willems, R. J., van Schaik, W., & Schürch, A. C. (2017). On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data. *Microbial Genomics*, *3*(10), e000128. https://doi.org/10.1099/mgen.0.000128

Arvin, E., Nielsen, L., Tully, A., Albrechtsen, H.-J., & Mosbæk, H. (2004). MTBE removal by biofiltration in a water works. In *2nd IWA leading-edge conference on water and wastewater treatment technologies*.

Asfaw, A., Maher, K., & Shucksmith, J. D. (2018). Modelling of metaldehyde concentrations in surface waters: A travel time based approach. *Journal of Hydrology*, *562*, 397–410. https://doi.org/10.1016/j.jhydrol.2018.04.074

Ashton, P. M., Nair, S., Dallman, T., Rubino, S., Rabsch, W., Mwaigwisya, S., et al. (2015). MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. *Nature Biotechnology*, *33*(3), 296–302. https://doi.org/10.1038/nbt.3103

Autin, O., Hart, J., Jarvis, P., MacAdam, J., Parsons, S. A., & Jefferson, B. (2012). Comparison of UV/H2O2 and UV/TiO2 for the degradation of metaldehyde: Kinetics and the impact of background organics. *Water Research*, *46*(17), 5655–5662. https://doi.org/10.1016/j.watres.2012.07.057

Autin, O., Hart, J., Jarvis, P., MacAdam, J., Parsons, S. A., & Jefferson, B. (2013a). Comparison of UV/TiO2 and UV/H2O2 processes in an annular photoreactor for removal of micropollutants: Influence of water parameters on metaldehyde removal, quantum yields and energy consumption. *Applied Catalysis B: Environmental*, *138*, 268–275. https://doi.org/10.1016/j.apcatb.2013.02.045

Autin, O., Hart, J., Jarvis, P., MacAdam, J., Parsons, S. A., & Jefferson, B. (2013b). The impact of background organic matter and alkalinity on the degradation of the pesticide metaldehyde by two advanced oxidation processes: UV/H2O2 and UV/TiO2. *Water Research*, *47*(6), 2041–2049. https://doi.org/10.1016/j.watres.2013.01.022

Autin, O., Romelot, C., Rust, L., Hart, J., Jarvis, P., MacAdam, J., et al. (2013). Evaluation of a UV-light emitting diodes unit for the removal of micropollutants in water for low energy advanced oxidation processes. *Chemosphere*, *92*(6), 745–751. https://doi.org/10.1016/j.chemosphere.2013.04.028

Bælum, J., & Jacobsen, C. S. (2009). TaqMan probe-based real-time PCR assay for detection and discrimination of class i, ii, and III tfdA genes in soils treated with phenoxy acid herbicides. *Applied and Environmental Microbiology*, *75*(9), 2969–2972. https://doi.org/10.1128/AEM.02051-08

Bælum, J., Nicolaisen, M. H., Holben, W. E., Strobel, B. W., Sørensen, J., & Jacobsen, C. S. (2008). Direct analysis of tfdA gene expression by indigenous bacteria in phenoxy acid amended agricultural soil. *ISME Journal*, *2*(6), 677–687. https://doi.org/10.1038/ismej.2008.21

Bælum, J., Prestat, E., David, M. M., Strobel, B. W., & Jacobsen, C. S. (2012). Modeling of phenoxy acid herbicide mineralization and growth of microbial degraders in 15 soils monitored by quantitative real-time PCR of the functional tfdA gene. *Applied and Environmental Microbiology*, *78*(15), 5305–5312. https://doi.org/10.1128/AEM.00990-12

Bai, Y., Liu, R., Liang, J., & Qu, J. (2013). Integrated Metagenomic and Physiochemical Analyses to Evaluate the Potential Role of Microbes in the Sand Filter of a Drinking Water Treatment System. *PLoS ONE*, *8*(4), e61011. https://doi.org/10.1371/journal.pone.0061011

Balashova, N., Wilderspin, S., Cai, C., & Reid, B. J. (2020). Ubiquity of microbial capacity to degrade metaldehyde in dissimilar agricultural, allotment and garden soils. *Science of the Total Environment*, *704*, 135412. https://doi.org/10.1016/j.scitotenv.2019.135412

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Bates, N., Sutton, N. M., & Campbell, A. (2012). Suspected metaldehyde slug bait poisoning in dogs: a retrospective analysis of cases reported to the Veterinary Poisons Information Service. *Veterinary Record*, *171*(13), 324. https://doi.org/10.1136/vr.100734

Ben Salem, A., Rouard, N., Devers, M., Béguet, J., Martin-Laurent, F., Caboni, P., et al. (2018). Environmental Fate of the Insecticide Chlorpyrifos in Soil Microcosms and Its Impact on Soil Microbial Communities. In *Recent advances in environmental science from the Euro-Mediterranean and surrounding regions: Proceedings of EuroMediterranean Conference for Environmental Integration (EMCEI-1), Tunisia 2017* (pp. 387–389). https://doi.org/10.1007/978-3-319-70548-4_122

Benner, J., Helbling, D. E., Kohler, H. P. E., Wittebol, J., Kaiser, E., Prasse, C., et al. (2013). Is biological treatment a viable alternative for micropollutant removal in drinking water

treatment processes? *Water Research*, *47*(16), 5955–5976. https://doi.org/10.1016/j.watres.2013.07.015

Bergthorsson, U., Andersson, D. I., & Roth, J. R. (2007). Ohno's dilemma: Evolution of new genes under continuous selection. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(43), 17004–17009. https://doi.org/10.1073/pnas.0707158104

Bers, K., Leroy, B., Breugelmans, P., Albers, P., Lavigne, R., Sørensen, S. R., et al. (2011). A novel hydrolase identified by genomic-proteomic analysis of phenylurea herbicide mineralization by Variovorax sp. strain SRS16. *Applied and Environmental Microbiology*, *77*(24), 8754–8764. https://doi.org/10.1128/AEM.06162-11

Beulke, S., Dubus, I. G., Brown, C. D., & Gottesbüren, B. (2000). Simulation of Pesticide Persistence in the Field on the Basis of Laboratory Data—A Review. *Journal of Environmental Quality*, *29*(5), 1371–1379. https://doi.org/10.2134/jeq2000.00472425002900050001x

Bieri, M. (2003). The environmental profile of metaldehyde. In *BCPC Symposium Proceedings* (pp. 255–262). British Crop Protection Council.

Blackwell, G. A., Nigro, S. J., & Hall, R. M. (2016). Evolution of AbGRI2-0, the progenitor of the AbGRI2 resistance island in global clone 2 of Acinetobacter baumannii. *Antimicrobial Agents and Chemotherapy*, *60*(3), 1421–1429. https://doi.org/10.1128/AAC.02662-15

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bond, S. G. (2018). *Assessing the fate of Metaldehyde Applied to Arable Soils*. University of Leicester.

Bosma, T. N. P., Middeldorp, P. J. M., Schraa, G., & Zehnder, A. J. B. (1996). Mass transfer limitation of biotransformation: Quantifying bioavailability. *Environmental Science and Technology*, *31*(1), 248–252. https://doi.org/10.1021/es960383u

Bouchez, T., Patureau, D., Dabert, P., Juretschko, S., Doré, J., Delgenès, P., et al. (2000). Ecological study of a bioaugmentation failure. *Environmental Microbiology*, *2*(2), 179–190. https://doi.org/10.1046/j.1462-2920.2000.00091.x

Bourne, D. G., Blakeley, R. L., Riddles, P., & Jones, G. J. (2006). Biodegradation of the cyanobacterial toxin microcystin LR in natural water and biologically active slow sand filters. *Water Research*, *40*(6), 1294–1302. https://doi.org/10.1016/j.watres.2006.01.022

Brady, J. A., Wallender, W. W., Werner, I., Fard, B. M., Zalom, F. G., Oliver, M. N., et al. (2006). Pesticide runoff from orchard floors in Davis, California, USA: A comparative analysis of diazinon and esfenvalerate. *Agriculture, Ecosystems and Environment*, *115*(1–4), 56–68. https://doi.org/10.1016/j.agee.2005.12.009

Busquets, R., Kozynchenko, O. P., Whitby, R. L. D., Tennison, S. R., & Cundy, A. B. (2014). Phenolic carbon tailored for the removal of polar organic contaminants from water: A solution to the metaldehyde problem? *Water Research*, *61*, 46–56. https://doi.org/10.1016/j.watres.2014.04.048

Bustin, S. (2000). Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology*, *25*(2), 169–193. https://doi.org/10.1677/jme.0.0250169

Bustin, S., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2009). The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, *55*(4), 611–622. https://doi.org/10.1373/clinchem.2008.112797

Cabezón, E., Ripoll-Rozada, J., Peña, A., de la Cruz, F., & Arechaga, I. (2015). Towards an integrated model of bacterial conjugation. *FEMS Microbiology Reviews*, *39*(1), 81–95. https://doi.org/10.1111/1574-6976.12085

Cakmakci, M., Koyuncu, I., & Kinaci, C. (2008). Effects of iron concentrations, filter hydraulic loading rates, and porosities on iron removal by rapid sand filtration. *Environmental Engineering Science*, *25*(5), 669–676. https://doi.org/10.1089/ees.2007.0060

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–3. https://doi.org/10.1038/nmeth.3869

Candela, L. (2003). El transporte de los plaguicidas a las aguas subterráneas. *Boletin Geologico y Minero*, *114*(4), 409–417.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, *7*(5), 335–336. https://doi.org/10.1038/nmeth.f.303

Carpenter, M. (1989). *Metaldehyde Draft Assessment Report , Vol. 3, B8*.

Carter, A. (2000). How pesticides get into water - And proposed reduction measures. *Pesticide Outlook*, *11*(4), 149–156. https://doi.org/10.1039/b006243j

Carvalhais, V., Delgado-Rastrollo, M., Melo, L. D. R., & Cerca, N. (2013). Controlled RNA contamination and degradation and its impact on qPCR gene expression in S. epidermidis biofilms. *Journal of Microbiological Methods*, *95*(2), 195–200. https://doi.org/10.1016/j.mimet.2013.08.010

Carver, T., Harris, S. R., Berriman, M., Parkhill, J., & Mcquillan, J. A. (2012). Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data, *28*(4), 464–469. https://doi.org/10.1093/bioinformatics/btr703

Carver, T., Thomson, N., Bleasby, A., Berriman, M., & Parkhill, J. (2009). DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics*, *25*(1), 119–120. https://doi.org/10.1093/bioinformatics/btn578

Castillo, J. M., Beguet, J., Martin-Laurent, F., & Romero, E. (2016). Multidisciplinary assessment of pesticide mitigation in soil amended with vermicomposted agroindustrial wastes. *Journal of Hazardous Materials*, *304*, 379–387. https://doi.org/10.1016/j.jhazmat.2015.10.056

Castle, G. D., Mills, G. A., Gravell, A., Jones, L., Townsend, I., Cameron, D. G., & Fones, G. R. (2017). Review of the molluscicide metaldehyde in the environment. *Environmental Science: Water Research & Technology*, *3*(3), 415–428. https://doi.org/10.1039/C7EW00039A

Castro-Gutiérrez, V., Fuller, E., Thomas, J. C., Sinclair, C. J., Johnson, S., Helgason, T., & Moir, J. W. B. (2020). Genomic basis for pesticide degradation revealed by selection, isolation and characterisation of a library of metaldehyde-degrading strains from soil. *Soil Biology and Biochemistry*, *142*, 107702. https://doi.org/10.1016/j.soilbio.2019.107702

Castro-Gutiérrez, V., Masís-Mora, M., Caminal, G., Vicent, T., Carazo-Rojas, E., Mora-López, M., & Rodríguez-Rodríguez, C. E. (2016). A microbial consortium from a biomixture swiftly degrades high concentrations of carbofuran in fluidized-bed reactors. *Process Biochemistry*, *51*(10), 1585–1593. https://doi.org/10.1016/J.PROCBIO.2016.07.003

Castro-Gutiérrez, V., Masís-Mora, M., Carazo-Rojas, E., Mora-López, M., & Rodríguez-Rodríguez, C. E. (2018). Impact of oxytetracycline and bacterial bioaugmentation on the efficiency and microbial community structure of a pesticide-degrading biomixture. *Environmental Science and Pollution Research*, *25*(12), 11787–11799. https://doi.org/10.1007/s11356-018-1436-1

Chandler, D. P. (1998). Redefining relativity: Quantitative PCR at low template concentrations for industrial and environmental microbiology. *Journal of Industrial Microbiology and Biotechnology*, *21*(3), 128–140. https://doi.org/10.1038/sj.jim.2900546

Changey, F., Devers-Lamrani, M., Rouard, N., & Martin-Laurent, F. (2011). In vitro evolution of an atrazine-degrading population under cyanuric acid selection pressure: Evidence for the selective loss of a 47kb region on the plasmid ADP1 containing the atzA, B and C genes. *Gene*, *490*(1–2), 18–25. https://doi.org/10.1016/j.gene.2011.09.005

Chiang, Y., & Kresge, A. J. (1985). Kinetics of hydrolysis of acetaldehyde ethyl hemiacetal in aqueous solution. *The Journal of Organic Chemistry*, *50*(25), 5038–5040. https://doi.org/10.1021/jo00225a007

Chief Inspector of Drinking Water. (2017a). *Drinking Water 2016- Summary of the Chief Inspector's Report for Drinking Water in England*. London.

Chief Inspector of Drinking Water. (2017b). *Drinking Water 2016 – Public water supplies for England and Wales, Quarter 3, July-September 2016*. London.

Chief Inspector of Drinking Water. (2018). *Drinking Water 2017 - Summary of the Chief Inspector's Report for Drinking Water in England*. London.

Chief Inspector of Drinking Water. (2019). *Drinking Water 2018 - Summary of the Chief Inspector's Report for Drinking Water in England*. London.

Cirja, M., Ivashechkin, P., Schäffer, A., & Corvini, P. F. X. (2008). Factors affecting the removal of organic micropollutants from wastewater in conventional treatment plants (CTP) and membrane bioreactors (MBR). *Reviews in Environmental Science and Biotechnology*, *7*(1), 61–78. https://doi.org/10.1007/s11157-007-9121-8

Clarke, K. R., Gorley, R., Somerfield, P., & Warwick, R. (2014). *Change in marine communities: an approach to statistical analysis and interpretation* (3rd ed.). Plymouth, UK: PRIMER-E.

Clasen, J. (1998). Efficiency control of particle removal by rapid sand filters in treatment plants fed with reservoir water: A survey of different methods. *Water Science and Technology*, *37*(2), 19–26. https://doi.org/10.1016/S0273-1223(98)00005-5

Collier, L. S., Gaines, G. L., & Neidle, E. L. (1998). Regulation of benzoate degradation in Acinetobacter sp. strain ADP1 by BenM, a LysR-type transcriptional activator. *Journal of Bacteriology*, *180*(9), 2493–2501. https://doi.org/10.1128/jb.180.9.2493-2501.1998

Collins, M. R. (1998). Experiences introducing '"new"' technology: slow sand filtration. In J. Cotruvo, G. Craun, & N. Hearn (Eds.), *Providing safe drinking water in small systems: technology, operations, and economics. Proceedings of the 1st International Symposium on Safe Drinking Water in Small Systems: technology, operations, and economics* (pp. 213–224). Boca Raton, Florida: CRC Press.

Cope, R. B., White, K. S., More, E., Holmes, K., Nair, A., Chauvin, P., & Oncken, A. (2006). Exposure-to-treatment interval and clinical severity in canine poisoning: A retrospective analysis at a Portland Veterinary Emergency Center. *Journal of Veterinary Pharmacology and Therapeutics*, *29*(3), 233–236. https://doi.org/10.1111/j.1365-2885.2006.00730.x

Copley, S. D. (2009). Evolution of efficient pathways for degradation of anthropogenic chemicals. *Nature Chemical Biology*, *5*(8), 559–566. https://doi.org/10.1038/nchembio.197

Copley, S. D. (2020). Evolution of new enzymes by gene duplication and divergence. *The FEBS Journal*, *287*(7), 1262–1283. https://doi.org/10.1111/febs.15299

Corfitzen, C. B., Albrechtsen, H. J., & Arvin, E. (2009). Removal of the phenoxyacide herbicide Mecoprop (MCPP) in water works rapid sandfilters. In *Water Quality Technology Conference (WQTC) and Exposition. Quality Water in a High-Tech Environment*. American Water Works Association (AWWA).

Cosgrove, S., Jefferson, B., & Jarvis, P. (2019). Pesticide removal from drinking water sources by adsorption: a review. *Environmental Technology Reviews*, *8*(1), 1–24. https://doi.org/10.1080/21622515.2019.1593514

Cragg, J. B., & Vincent, M. H. (1952). The action of metaldehyde on the slug Agrolimax reticulatus (Müller). *Annals of Applied Biology*, *39*(3), 392–406. https://doi.org/10.1111/j.1744-7348.1952.tb01023.x

Cranor, W. (1990). *Metaldehyde Draft Assessment Report, Vol. 3, B8*.

Crouzet, O., Batisson, I., Besse-Hoggan, P., Bonnemoy, F., Bardot, C., Poly, F., et al. (2010). Response of soil microbial communities to the herbicide mesotrione: A dose-effect microcosm approach. *Soil Biology and Biochemistry*, *42*(2), 193–202. https://doi.org/10.1016/J.SOILBIO.2009.10.016

Cycoń, M., Markowicz, A., Borymski, S., Wójcik, M., & Piotrowska-Seget, Z. (2013). Imidacloprid induces changes in the structure, genetic diversity and catabolic activity of soil microbial communities. *Journal of Environmental Management*, *131*, 55–65. https://doi.org/10.1016/J.JENVMAN.2013.09.041

D'Alessio, M., Yoneyama, B., Kirs, M., Kisand, V., & Ray, C. (2015). Pharmaceutically active compounds: Their removal during slow sand filtration and their impact on slow sand filtration bacterial removal. *Science of the Total Environment*, *524–525*, 124–135. https://doi.org/10.1016/j.scitotenv.2015.04.014

De Roy, K., Marzorati, M., Negroni, A., Thas, O., Balloi, A., Fava, F., et al. (2013). Environmental conditions and community evenness determine the outcome of biological invasion. *Nature Communications*, *4*(1), 1–5. https://doi.org/10.1038/ncomms2392

De Schrijver, A., & De Mot, R. (1999). Degradation of pesticides by actinomycetes. *Critical Reviews in Microbiology*, *25*(2), 85–119. https://doi.org/10.1080/10408419991299194

De Souza, M. L., Newcombe, D., Alvey, S., Crowley, D. E., Hay, A., Sadowsky, M. J., & Wackett, L. P. (1998). Molecular basis of a bacterial consortium: Interspecies catabolism of atrazine. *Applied and Environmental Microbiology*, *64*(1), 168–174. https://doi.org/10.1128/aem.64.1.178-184.1998

DEFRA. (2018). Restrictions on the use of metaldehyde to protect wildlife. https://www.gov.uk/government/news/restrictions-on-the-use-of-metaldehyde-to-protect-wildlife. Accessed 6 June 2020

Dejonghe, W., Goris, J., Fantroussi, S. E., Hofte, M., Vos, P. D., Verstraete, W., & Top, E. M. (2000). Effect of dissemination of 2,4-dichlorophenoxyacetic acid (2,4-D) degradation plasmids on 2,4-d degradation and on bacterial community structure in two different soil horizons. *Applied and Environmental Microbiology*, *66*(8), 3297–3304. https://doi.org/10.1128/AEM.66.8.3297-3304.2000

Delmont, T. O., Simonet, P., & Vogel, T. M. (2012). Describing microbial communities and performing global comparisons in the omic era. *ISME Journal*, *6*(9), 1625–1628. https://doi.org/10.1038/ismej.2012.55

Dennis, J. J. (2005). The evolution of IncP catabolic plasmids. *Current Opinion in Biotechnology*, *16*(3), 291–298. https://doi.org/10.1016/j.copbio.2005.04.002

Diaz-Aroca, E., Mendiola, M. V., Zabala, J. C., & De La Cruz, F. (1987). Transposition of IS91 does not generate a target duplication. *Journal of Bacteriology*, *169*(1), 442–443. https://doi.org/10.1128/jb.169.1.442-443.1987

Diez, M. C. (2010). Biological aspects involved in the degradation of organic pollutants. *Journal of Soil Science and Plant Nutrition*, *10*(3), 244–267. https://doi.org/10.4067/S0718-95162010000100004

DiGiovanni, G. D., Neilson, J. W., Pepper, I. L., & Sinclair, N. A. (1996). Gene transfer of Alcaligenes eutrophus JMP134 plasmid pJP4 to indigenous soil recipients. *Applied and environmental microbiology*, *62*(7), 2521–6.

Dillon, G., Hall, T., Jönsson, J., Rockett, L., Shepherd, D., Watts, M., & Webber, F. (2013). *Emerging pesticides; what next? Report Ref. No. 13/DW/14/6*.

Dinamarca, M. A., Cereceda-Balic, F., Fadic, X., & Seeger, M. (2007). Analysis of s-triazine-degrading microbial communities in soils using most-probable-number enumeration and tetrazolium-salt detection. *International Microbiology*, *10*(3), 209–15.

Dobritsa, A. P., & Samadpour, M. (2016). Transfer of eleven species of the genus Burkholderia to the genus Paraburkholderia and proposal of Caballeronia gen. nov. to accommodate twelve species of the genera Burkholderia and Paraburkholderia. *International Journal of Systematic and Evolutionary Microbiology*, *66*(8), 2836–2846. https://doi.org/10.1099/ijsem.0.001065

Dolan, T., Howsam, P., & Parsons, D. J. (2012). Diffuse pesticide pollution of drinking water sources: Impact of legislation and UK responses. *Water Policy*, *14*(4), 680–693. https://doi.org/10.2166/wp.2012.147

Doria, F. C., Borges, A. C., Kim, J. K., Nathan, A., Joo, J. C., & Campos, L. C. (2013). Removal of metaldehyde through photocatalytic reactions using nano-sized zinc oxide composites. *Water, Air, and Soil Pollution*, *224*, 1434. https://doi.org/10.1007/s11270-013-1434-3

Eckert, M., Fleischmann, G., Reinhard, J., Bolt, H., & Golka, K. (2012). Acetaldehyde. In *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley-VCH Verlag GmbH & Co.

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Egli, T. (2010). How to live at very low substrate concentration. *Water Research*, *44*(17), 4826–4837. https://doi.org/10.1016/j.watres.2010.07.023

Ekblom, R., & Wolf, J. B. W. (2014). A field guide to whole-genome sequencing, assembly and

annotation. *Evolutionary Applications*, *7*(9), 1026–1042. https://doi.org/10.1111/eva.12178

Ellehnorn, M., & Barceloux, D. (1997). *Ellenhorn's medical toxicology: Diagnosis and treatment of human poisoning* (2nd ed.). Philadelphia. https://doi.org/10.1016/s0736-4679(97)00099-1

Escolà Casas, M., & Bester, K. (2015). Can those organic micro-pollutants that are recalcitrant in activated sludge treatment be removed from wastewater by biofilm reactors (slow sand filters)? *Science of the Total Environment*, *506–507*, 315–322. https://doi.org/10.1016/j.scitotenv.2014.10.113

European Food Safety Authority. (2010a). Conclusion on the peer review of the pesticide risk assessment of the active substance metaldehyde. *EFSA Journal*, *8*(10), 1856. https://doi.org/10.2903/j.efsa.2010.1856

European Food Safety Authority. (2010b). Conclusion on the peer review of the pesticide risk assessment of the active substance metaldehyde. *EFSA Journal*, *8*(10), 1–71. https://doi.org/10.2903/j.efsa.2010.1856

Fang, H., Zhang, H., Han, L., Mei, J., Ge, Q., Long, Z., & Yu, Y. (2018). Exploring bacterial communities and biodegradation genes in activated sludge from pesticide wastewater treatment plants via metagenomic analysis. *Environmental Pollution*, *243*, 1206–1216. https://doi.org/10.1016/j.envpol.2018.09.080

Fenner, K., Canonica, S., Wackett, L. P., & Elsner, M. (2013). Evaluating pesticide degradation in the environment: Blind spots and emerging opportunities. *Science*, *341*(6147), 752–758. https://doi.org/10.1126/science.1236281

Fera Science. (2020). PUS Stats Home. http://pusstats.fera.defra.gov.uk/myresults.cfm. Accessed 6 June 2020

Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology*, *15*(10), 579–590. https://doi.org/10.1038/nrmicro.2017.87

Fitzpatrick, C. S., & Gregory, J. (2003). Coagulation and filtration. In D. Mara & N. Horan (Eds.), *Handbook of Water and Wastewater Microbiology*. Elsevier. https://doi.org/10.1016/B978-012470100-7/50039-X

FOCUS. (2006). *Guidance Document on Estimating Persistence and Degradation Kinetics from Environmental Fate Studies on Pesticides in EU Registration*.

Forbes, J. D., Knox, N. C., Ronholm, J., Pagotto, F., & Reimer, A. (2017). Metagenomics: The next culture-independent game changer. *Frontiers in Microbiology*, *8*, 1069. https://doi.org/10.3389/fmicb.2017.01069

Gaby, J. C., & Buckley, D. H. (2017). The Use of Degenerate Primers in qPCR Analysis of Functional Genes Can Cause Dramatic Quantification Bias as Revealed by Investigation of nifH Primer Performance. *Microbial Ecology*, *74*(3), 701–708. https://doi.org/10.1007/s00248-017-0968-0

Galiulin, R. V., Bashkin, V. N., Galiulina, R. A., & Birch, P. (2001). The theoretical basis of microbiological transformation and degradation of pesticides in soil. *Land Contamination and Reclamation*, *9*(4), 367–376.

Gallego, S., Devers-Lamrani, M., Rousidou, K., Karpouzas, D. G., & Martin-Laurent, F. (2019).

Assessment of the effects of oxamyl on the bacterial community of an agricultural soil exhibiting enhanced biodegradation. *Science of the Total Environment*, *651*, 1189–1198. https://doi.org/10.1016/j.scitotenv.2018.09.255

García, J. (1997). *Introducción a los plaguicidas*. San José: Editorial Universidad Estatal a Distancia.

Garcillán-Barcia, M., Bernales, I., Mendiola, M. V., & Cruz, F. D. La. (2002). IS91 Rolling-Circle Transposition. In N. Craig (Ed.), *Mobile DNA II* (pp. 891–904). Washington D.C.: American Society of Microbiology. https://doi.org/10.1128/9781555817954.ch37

Garcillán-Barcia, M., Bernales, I., Victoria Mendiola, M., & De la Cruz, F. (2001). Single-stranded DNA intermediates in IS91 rolling-circle transposition. *Molecular Microbiology*, *39*(2), 494–502. https://doi.org/10.1046/j.1365-2958.2001.02261.x

Garcillán-Barcia, M., & de la Cruz, F. (2002). Distribution of IS91 family insertion sequences in bacterial genomes: evolutionary implications. *FEMS Microbiology Ecology*, *42*(2), 303–313. https://doi.org/10.1111/j.1574-6941.2002.tb01020.x

Garthwaite, D., Barker, I., Ridley, L., Mace, A., Parrish, G., MacArthur, R., & Lu, Y. (2018). *Pesticide Usage Survey Report 271 - Arable crops in the United Kingdom (Version 2) 2016*. Sand Hutton.

George, S., Pankhurst, L., Hubbard, A., Votintseva, A., Stoesser, N., Sheppard, A. E., et al. (2017). Resolving plasmid structures in enterobacteriaceae using the MinION nanopore sequencer: Assessment of MinION and MinION/illumina hybrid data assembly approaches. *Microbial Genomics*, *3*(8). https://doi.org/10.1099/mgen.0.000118

Geyikci, F. (2011). Pesticides and Their Movement Surface Water and Ground Water. In M. Stoytcheva (Ed.), *Pesticides in the Modern World - Risks and Benefits* (pp. 411–422). https://doi.org/10.5772/17301

Gianfreda, L., & Rao, M. A. (2004). Potential of extra cellular enzymes in remediation of polluted soils: A review. *Enzyme and Microbial Technology*, *35*(4), 339–354. https://doi.org/10.1016/j.enzmictec.2004.05.006

Gimingham, C. T. (1940). Some recent contributions by english workers to the development of methods of insect control. *Annals of Applied Biology*, *27*(2), 161–175. https://doi.org/10.1111/j.1744-7348.1940.tb07486.x

Goda, S. K., Elsayed, I. E., Khodair, T. A., El-Sayed, W., & Mohamed, M. E. (2010). Screening for and isolation and identification of malathion-degrading bacteria: cloning and sequencing a gene that potentially encodes the malathion-degrading enzyme, carboxylestrase in soil bacteria. *Biodegradation*, *21*(6), 903–913. https://doi.org/10.1007/s10532-010-9350-3

Goldstein, R. M., Mallory, L. M., & Alexander, M. (1985). Reasons for possible failure of inoculation to enhance biodegradation. *Applied and Environmental Microbiology*, *50*(4), 977–983. https://doi.org/10.1128/aem.50.4.977-983.1985

Goldstein, S., Beka, L., Graf, J., & Klassen, J. L. (2019). Evaluation of strategies for the assembly of diverse bacterial genomes using MinION long-read sequencing. *BMC Genomics*, *20*(1), 23. https://doi.org/10.1186/s12864-018-5381-7

Golosova, O., Henderson, R., Vaskin, Y., Gabrielian, A., Grekhov, G., Nagarajan, V., et al. (2014). Unipro UGENE NGS pipelines and components for variant calling, RNA-seq and ChIP-seq data analyses. *PeerJ*, *2*, e644. https://doi.org/10.7717/peerj.644

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, *17*(6), 333–351. https://doi.org/10.1038/nrg.2016.49

Gorski, L., Rivadeneira, P., & Cooley, M. B. (2019). New strategies for the enumeration of enteric pathogens in water. *Environmental Microbiology Reports*, *11*(6), 765–776. https://doi.org/10.1111/1758-2229.12786

Gupta, R. C. (2018). *Veterinary Toxicology: Basic and Clinical Principles* (3rd ed.). San Diego: Academic Press.

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, *29*(8), 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Hacker, J., & Carniel, E. (2001). Ecological fitness, genomic islands and bacterial pathogenicity. *EMBO reports*, *2*(5), 376–381. https://doi.org/10.1093/embo-reports/kve097

Haig, S., Collins, G., Davies, R. L., Dorea, C. C., & Quince, C. (2011). Biological aspects of slow sand filtration: Past, present and future. *Water Science and Technology: Water Supply*, *11*(4), 468–472. https://doi.org/10.2166/ws.2011.076

Haig, S., Gauchotte-Lindsay, C., Collins, G., & Quince, C. (2016). Bioaugmentation Mitigates the Impact of Estrogen on Coliform-Grazing Protozoa in Slow Sand Filters. *Environmental Science and Technology*, *50*(6), 3101–3110. https://doi.org/10.1021/acs.est.5b05027

Haig, S., Quince, C., Davies, R. L., Dorea, C. C., & Collins, G. (2014). Replicating the microbial community and water quality performance of full-scale slow sand filters in laboratory-scale filters. *Water Research*, *61*, 141–151. https://doi.org/10.1016/j.watres.2014.05.008

Haig, S., Quince, C., Davies, R. L., Dorea, C. C., & Collinsa, G. (2015). The relationship between microbial community evenness and function in slow sand filters. *mBio*, *6*(5), e00729-15. https://doi.org/10.1128/mBio.00729-15

Harmer, C. J., & Hall, R. M. (2016). IS 26 -Mediated Formation of Transposons Carrying Antibiotic Resistance Genes . *mSphere*, *1*(2). https://doi.org/10.1128/msphere.00038-16

Harmer, C. J., & Hall, R. M. (2019). An analysis of the IS6/IS26 family of insertion sequences: Is it a single family? *Microbial Genomics*, *5*(9). https://doi.org/10.1099/mgen.0.000291

Harmer, C. J., & Hall, R. M. (2020). IS 26 Family Members IS 257 and IS 1216 Also Form Cointegrates by Copy-In and Targeted Conservative Routes . *mSphere*, *5*(1). https://doi.org/10.1128/msphere.00811-19

Harmer, C. J., Moran, R. A., & Hall, R. M. (2014). Movement of IS26-Associated antibiotic resistance genes occurs via a translocatable unit that includes a single IS26 and preferentially inserts adjacent to another IS26. *mBio*, *5*(5), e01801-14. https://doi.org/10.1128/mBio.01801-14

Harms, H., Schlosser, D., & Wick, L. Y. (2011). Untapped potential: Exploiting fungi in bioremediation of hazardous chemicals. *Nature Reviews Microbiology*, *9*(3), 177–192. https://doi.org/10.1038/nrmicro2519

Hayden, R. T., Hokanson, K. M., Pounds, S. B., Bankowski, M. J., Belzer, S. W., Carr, J., et al. (2008). Multicenter comparison of different real-time PCR assays for quantitative detection of Epstein-Barr virus. *Journal of Clinical Microbiology*, *46*(1), 157–163. https://doi.org/10.1128/JCM.01252-07

He, P., & Moran, G. R. (2011). Structural and mechanistic comparisons of the metal-binding members of the vicinal oxygen chelate (VOC) superfamily. *Journal of Inorganic Biochemistry*, *105*(10), 1259–1272. https://doi.org/10.1016/J.JINORGBIO.2011.06.006

Heid, C. A., Stevens, J., Livak, K. J., & Williams, P. M. (1996). Real time quantitative PCR. *Genome Research*, *6*(10), 986–994. https://doi.org/10.1101/gr.6.10.986

Helbling, D. E., Hollender, J., Kohler, H. P. E., Singer, H., & Fenner, K. (2010). High-throughput identification of microbial transformation products of organic micropollutants. *Environmental Science and Technology*, *44*(17), 6621–6627. https://doi.org/10.1021/es100970m

Hiraishi, A., & Ueda, Y. (1994). Rhodoplanes gen. nov., a new genus of phototrophic bacteria including Rhodopseudomonas rosea as Rhodoplanes roseus comb. nov. and Rhodoplanes elegans sp. nov. *International Journal of Systematic Bacteriology*, *44*(4), 665–673. https://doi.org/10.1099/00207713-44-4-665

Ho, L., Hoefel, D., Bock, F., Saint, C. P., & Newcombe, G. (2007). Biodegradation rates of 2-methylisoborneol (MIB) and geosmin through sand filters and in bioreactors. *Chemosphere*, *66*(11), 2210–2218. https://doi.org/10.1016/j.chemosphere.2006.08.016

Holben, W. E., Schroeter, B. M., Calabrese, V. G. M., Olsen, R. H., Kukor, J. K., Biederbeck, V. O., et al. (1992). Gene probe analysis of soil microbial populations selected by amendment with 2,4-dichlorophenoxyacetic acid. *Applied and Environmental Microbiology*, *58*(12), 3941–3948.

Horemans, B., Vandermaesen, J., Vanhaecke, L., Smolders, E., & Springael, D. (2013). Variovorax sp.-mediated biodegradation of the phenyl urea herbicide linuron at micropollutant concentrations and effects of natural dissolved organic matter as supplementary carbon source. *Applied Microbiology and Biotechnology*, *97*(22), 9837–9846. https://doi.org/10.1007/s00253-013-4690-7

Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., Von Mering, C., & Bork, P. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Molecular Biology and Evolution*, *34*(8), 2115–2122. https://doi.org/10.1093/molbev/msx148

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*, *47*(D1), D309–D314. https://doi.org/10.1093/nar/gky1085

Iida, S., Mollet, B., Meyer, J., & Arber, W. (1984). Functional characterization of the prokaryotic mobile genetic element IS 26. *MGG Molecular & General Genetics*, *198*(1), 84–89. https://doi.org/10.1007/BF00328705

Itoh, H., Navarro, R., Takeshita, K., Tago, K., Hayatsu, M., Hori, T., & Kikuchi, Y. (2014). Bacterial population succession and adaptation affected by insecticide application and soil spraying history. *Frontiers in Microbiology*, *5*, 457. https://doi.org/10.3389/fmicb.2014.00457

Jaworski, J. P., Pluta, A., Rola-Luszczak, M., McGowan, S. L., Finnegan, C., Heenemann, K., et al. (2018). Interlaboratory comparison of six real-time PCR assays for detection of bovine leukemia virus proviral DNA. *Journal of Clinical Microbiology*, *56*(7), e00304-18. https://doi.org/10.1128/JCM.00304-18

Jeffries, T. C., Rayu, S., Nielsen, U. N., Lai, K., Ijaz, A., Nazaries, L., & Singh, B. (2018).

Metagenomic functional potential predicts degradation rates of a model organophosphorus xenobiotic in pesticide contaminated soils. *Frontiers in Microbiology*, *9*, 1–12. https://doi.org/10.3389/fmicb.2018.00147

Jezbera, J., Horňák, K., & Šimek, K. (2006). Prey selectivity of bacterivorous protists in different size fractions of reservoir water amended with nutrients. *Environmental Microbiology*, *8*(8), 1330–1339. https://doi.org/10.1111/j.1462-2920.2006.01026.x

Jin, Y. O., & Mattes, T. E. (2011). Assessment and modification of degenerate qPCR primers that amplify functional genes from etheneotrophs and vinyl chloride-assimilators. *Letters in Applied Microbiology*, *53*(5), 576–580. https://doi.org/10.1111/j.1472-765X.2011.03144.x

Joernvall, H., & Persson, B. (2006). Aldehyde dehydrogenases active sites. https://prosite.expasy.org/PDOC00068. Accessed 16 April 2020

Johnson, G. R., & Spain, J. C. (2003). Evolution of catabolic pathways for synthetic compounds: Bacterial pathways for degradation of 2,4-dinitrotoluene and nitrobenzene. *Applied Microbiology and Biotechnology*, *62*(2–3), 110–123. https://doi.org/10.1007/s00253-003-1341-4

Kalisky, T., Dekel, E., & Alon, U. (2007). Cost-benefit theory and optimal design of gene regulation functions. *Physical Biology*, *44*(4), 229–245. https://doi.org/10.1088/1478-3975/4/4/001

Kanitkar, Y. H., Stedtfeld, R. D., Hatzinger, P. B., Hashsham, S. A., & Cupples, A. M. (2017). Most probable number with visual based LAMP for the quantification of reductive dehalogenase genes in groundwater samples. *Journal of Microbiological Methods*, *143*, 44–49. https://doi.org/10.1016/j.mimet.2017.10.003

Kay, P., & Grayson, R. (2014). Using water industry data to assess the metaldehyde pollution problem. *Water and Environment Journal*, *28*(3), 410–417. https://doi.org/10.1111/wej.12056

Kegley, S. E., Hill, B. R., Orme, S., & Choi, A. H. (2016). PAN Pesticide Database, Pesticide Action Network, North America. http://www.pesticideinfo.org/. Accessed 30 April 2018

Kekulé, A., & Zincke, T. (1872). Ueber das sogenannte Chloraceten und die polymeren Modificationen des Aldehyds. *Justus Liebigs Annalen der Chemie*, *162*(1), 125–150. https://doi.org/10.1002/jlac.18721620111

Kem, J. (1996, January). *Full-scale comparative evaluation of two slow sand filter cleaning methods*. University of New Hampshire.

Kerle, E. a., Jenkins, J. J., & Vogue, P. a. (2007). Understanding pesticide persistence and mobility for groundwater and surface protection (http://extension. oregonstate. edu/catalog/html/em/em8561-e/, verified 05/2010). *Extention Service, Oregon State University*.

Kim, J., Lee, A. H., & Chang, W. (2018). Enhanced bioremediation of nutrient-amended, petroleum hydrocarbon-contaminated soils over a cold-climate winter. *Science of the Total Environment*, *612*, 903–913. https://doi.org/10.1016/j.scitotenv.2017.08.227

Kim, M., Oh, H.-S., Park, S.-C., & Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, *64*, 346–351. https://doi.org/10.1099/ijs.0.059774-0

Kingsford, C., Schatz, M. C., & Pop, M. (2010). Assembly complexity of prokaryotic genomes

using short reads. *BMC Bioinformatics*, *11*(1), 21.

Klein, D. (2002). Quantification using real-time PCR technology: Applications and limitations. *Trends in Molecular Medicine*, *8*(6), 257–260. https://doi.org/10.1016/S1471-4914(02)02355-9

Kolvenbach, B. A., Helbling, D. E., Kohler, H. P. E., & Corvini, P. F. X. (2014). Emerging chemicals and the evolution of biodegradation capacities and pathways in bacteria. *Current Opinion in Biotechnology*, *27*, 8–14. https://doi.org/10.1016/j.copbio.2013.08.017

Krutz, J., Shaner, D. L., Weaver, M. A., Webb, R. M. T., Zablotowicz, R. M., Reddy, K. N., et al. (2010). Agronomic and environmental implications of enhanced s-triazine degradation. *Pest Management Science*, *66*(5), 461–481. https://doi.org/10.1002/ps.1909

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular biology and evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Kundu, K., Marozava, S., Ehrl, B., Merl-Pham, J., Griebler, C., & Elsner, M. (2019). Defining lower limits of biodegradation: atrazine degradation regulated by mass transfer and maintenance demand in Arthrobacter aurescens TC1. *ISME Journal*, *13*(9), 2236–2251. https://doi.org/10.1038/s41396-019-0430-z

Kurm, V., Van Der Putten, W. H., De Boer, W., Naus-Wiezer, S., & Gera Hol, W. H. (2017). Low abundant soil bacteria can be metabolically versatile and fast growing. *Ecology*, *98*(2), 555–564. https://doi.org/10.1002/ecy.1670

Kuzyakov, Y., & Blagodatskaya, E. (2015). Microbial hotspots and hot moments in soil: Concept & review. *Soil Biology and Biochemistry*, *83*, 184–199. https://doi.org/10.1016/J.SOILBIO.2015.01.025

Kwak, Y., Kim, S. J., Rhee, I. K., & Shin, J. H. (2012). Application of quantitative real-time polymerase chain reaction on the assessment of organophosphorus compound degradation in in situ soil. *Journal of the Korean Society for Applied Biological Chemistry*, *55*(6), 757–763. https://doi.org/10.1007/s13765-012-2168-4

Lagesen, K., Hallin, P., Rødland, A., Staerfeldt, H.-H., Rognes, T., & Ussery, D. W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research*, *35*(9), 3100–3108. https://doi.org/10.1093/nar/gkm160

Lal, D., Jindal, S., Kumari, H., Jit, S., Nigam, A., Sharma, P., et al. (2015). Bacterial diversity and real-time PCR based assessment of linA and linB gene distribution at hexachlorocyclohexane contaminated sites. *Journal of Basic Microbiology*, *55*(3), 363–373. https://doi.org/10.1002/jobm.201300211

Lapworth, D. J., Baran, N., Stuart, M. E., & Ward, R. S. (2012). Emerging organic contaminants in groundwater: A review of sources, fate and occurrence. *Environmental Pollution*, *163*, 287–303. https://doi.org/10.1016/j.envpol.2011.12.034

Lazartigues, A., Thomas, M., Cren-Olivé, C., Brun-Bellut, J., Le Roux, Y., Banas, D., & Feidt, C. (2013). Pesticide pressure and fish farming in barrage pond in Northeastern France. Part II: residues of 13 pesticides in water, sediments, edible fish and their relationships. *Environmental Science and Pollution Research*, *20*(1), 117–125. https://doi.org/10.1007/s11356-012-1167-7

Lee, H., Gurtowski, J., Yoo, S., Nattestad, M., Marcus, S., Goodwin, S., et al. (2016). Third-generation sequencing and the future of genomics. *bioRxiv*, 048603.

https://doi.org/10.1101/048603

Leu, C., Singer, H., Stamm, C., Müller, S. R., & Schwarzenbach, R. P. (2004). Variability of herbicide losses from 13 fields to surface water within a small catchment after a controlled herbicide application. *Environmental Science and Technology*, *38*(14), 3835–3841. https://doi.org/10.1021/es0499593

Lewis, K. A., Tzilivakis, J., Warner, D. J., & Green, A. (2016). An international database for pesticide risk assessments and management. *Human and Ecological Risk Assessment: An International Journal*, *22*(4), 1050–1064. https://doi.org/10.1080/10807039.2015.1133242

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, X., Xie, Y., Liu, M., Tai, C., Sun, J., Deng, Z., & Ou, H.-Y. (2018). oriTfinder: a web-based tool for the identification of origin of transfers in DNA sequences of bacterial mobile genetic elements. *Nucleic Acids Research*, *46*, 229–234. https://doi.org/10.1093/nar/gky352

Liu, J., Zheng, Y., Lin, H., Wang, X., Li, M., Liu, Y., et al. (2019). Proliferation of hydrocarbon-degrading microbes at the bottom of the Mariana Trench. *Microbiome*, *7*(1), 1–13. https://doi.org/10.1186/s40168-019-0652-3

Liu, Q., Tang, J., Liu, X., Song, B., Zhen, M., & Ashbolt, N. J. (2017). Response of microbial community and catabolic genes to simulated petroleum hydrocarbon spills in soils/sediments from different geographic locations. *Journal of Applied Microbiology*, *123*(4), 875–885. https://doi.org/10.1111/jam.13549

Lowe, T. M., & Chan, P. P. (2016). tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Research*, *44*(W1), W54–W57. https://doi.org/10.1093/nar/gkw413

Luo, H., & Gao, F. (2019). DoriC 10.0: an updated database of replication origins in prokaryotic genomes including chromosomes and plasmids. *Nucleic Acids Research*, *47*(D1), D74–D77. https://doi.org/10.1093/nar/gky1014

Maheshwari, M., Abulreesh, H. H., Khan, M. S., Ahmad, I., & Pichtel, J. (2017). Horizontal Gene Transfer in Soil and the Rhizosphere: Impact on Ecological Fitness of Bacteria. In V. Meena, P. Mishra, J. Bisht, & A. Pattanayak (Eds.), *Agriculturally Important Microbes for Sustainable Agriculture* (pp. 111–130). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-10-5589-8_6

Mahillon, J., & Chandler, M. (1998). Insertion Sequences. *Microbiology and Molecular Biology Reviews*, *62*(3), 725–774. https://doi.org/10.1128/MMBR.62.3.725-774.1998

Mammeri, H., Poirel, L., Mangeney, N., & Nordmann, P. (2003). Chromosomal integration of a cephalosporinase gene from Acinetobacter baumannii into Oligella urethralis as a source of acquired resistance to β-lactams. *Antimicrobial Agents and Chemotherapy*, *47*(5), 1536–1542. https://doi.org/10.1128/AAC.47.5.1536-1542.2003

Mandelbaum, R. T., Wackett, L. P., & Allan, D. L. (1993). Mineralization of the s-triazine ring of atrazine by stable bacterial mixed cultures. *Applied and Environmental Microbiology*, *59*(6), 1695–1701. https://doi.org/10.1128/aem.59.6.1695-1701.1993

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., et al. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Research*, *43*(D1), D222–D226.

https://doi.org/10.1093/nar/gku1221

Martin-Laurent, F., Cornet, L., Ranjard, L., López-Gutiérrez, J. C., Philippot, L., Schwartz, C., et al. (2004). Estimation of atrazine-degrading genetic potential and activity in three French agricultural soils. *FEMS Microbiology Ecology*, *48*(3), 425–435. https://doi.org/10.1016/j.femsec.2004.03.008

Martin-Laurent, F., Piutti, S., Hallet, S., Wagschal, I., Philippot, L., Catroux, G., & Soulas, G. (2003). Monitoring of atrazine treatment on soil bacterial, fungal and atrazine-degrading communities by quantitative competitive PCR. *Pest Management Science*, *59*(3), 259–268. https://doi.org/10.1002/ps.630

McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., DeSantis, T. Z., Probst, A., et al. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *The ISME Journal*, *6*(3), 610–618. https://doi.org/10.1038/ismej.2011.139

McDowall, B., Hoefel, D., Newcombe, G., Saint, C. P., & Ho, L. (2009). Enhancing the biofiltration of geosmin by seeding sand filter columns with a consortium of geosmin-degrading bacteria. *Water Research*, *43*(2), 433–440. https://doi.org/10.1016/j.watres.2008.10.044

Mendiola, M. V., Bernales, I., & De La Cruz, F. (1994). Differential roles of the transposon termini in IS91 transposition. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(5), 1922–1926. https://doi.org/10.1073/pnas.91.5.1922

Metaldehyde Stewardship Group. (2020). MSG stewardship guidelines. https://www.getpelletwise.co.uk/home/msg-guidelines/. Accessed 17 June 2020

Meynet, P., Head, I. M., Werner, D., & Davenport, R. J. (2015). Re-evaluation of dioxygenase gene phylogeny for the development and validation of a quantitative assay for environmental aromatic hydrocarbon degraders. *FEMS Microbiology Ecology*, *91*(6), fiv049. https://doi.org/10.1093/femsec/fiv049

Miller, R. (1928). Poisoning by “Meta Fuel” Tablets (Metacetaldehyde). *Archives of Disease in Childhood*, *3*(18), 292. https://doi.org/10.1136/adc.3.18.292

Mindlin, S., Petrenko, A., Kurakov, A., Beletsky, A., Mardanov, A., & Petrova, M. (2016). Resistance of Permafrost and Modern Acinetobacter lwoffii Strains to Heavy Metals and Arsenic Revealed by Genome Analysis. *BioMed Research International*, 3970831. https://doi.org/10.1155/2016/3970831

Mohamad Ibrahim, I. H., Gilfoyle, L., Reynolds, R., & Voulvoulis, N. (2019). Integrated catchment management for reducing pesticide levels in water: Engaging with stakeholders in East Anglia to tackle metaldehyde. *Science of the Total Environment*, *656*, 1436–1447. https://doi.org/10.1016/j.scitotenv.2018.11.260

Möllerfeld, J., Römbke, J., & Heller, M. (1993). *Metaldehyde Draft Assessment Report, Vol. 3, B8*.

Mollet, B., Iida, S., & Arber, W. (1985). Gene organization and target specificity of the prokaryotic mobile genetic element IS 26. *MGG Molecular & General Genetics*, *201*(2), 198–203. https://doi.org/10.1007/BF00425660

Mollet, B., Iida, S., Shepherd, J., & Arber, W. (1983). Nucleotide sequence of IS26, a new prokaryotic mobile genetic etement. *Nucleic Acids Research*, *11*(18), 6319–6330. https://doi.org/10.1093/nar/11.18.6319

Monard, C., Martin-Laurent, F., Devers-Lamrani, M., Lima, O., Vandenkoornhuyse, P., & Binet, F.

(2010). Atz gene expressions during atrazine degradation in the soil drilosphere. *Molecular Ecology*, *19*(4), 749–759. https://doi.org/10.1111/j.1365-294X.2009.04503.x

Monard, C., Martin-Laurent, F., Lima, O., Devers-Lamrani, M., & Binet, F. (2013). Estimating the biodegradation of pesticide in soils by monitoring pesticide-degrading gene expression. *Biodegradation*, *24*(2), 203–213. https://doi.org/10.1007/s10532-012-9574-5

Morales, M. E., Allegrini, M., Basualdo, J., Villamil, M. B., & Zabaloy, M. C. (2020). Primer design to assess bacterial degradation of glyphosate and other phosphonates. *Journal of Microbiological Methods*, *169*, 105814. https://doi.org/10.1016/j.mimet.2019.105814

Moran, R. A., & Hall, R. M. (2019). pBuzz: A cryptic rolling-circle plasmid from a commensal Escherichia coli has two inversely oriented oriTs and is mobilised by a B/O plasmid. *Plasmid*, *101*, 10–19. https://doi.org/10.1016/j.plasmid.2018.11.001

Morgenthaler, A., Kinney, W., Ebmeier, C., Walsh, C., Snyder, D., Cooper, V., et al. (2019). Mutations that improve the efficiency of a weak-link enzyme are rare compared to adaptive mutations elsewhere in the genome. *eLife*, *8*, e53535. https://doi.org/10.1101/624205

Mukherjee, N., Bartelli, D., Patra, C., Chauhan, B. V., Dowd, S. E., & Banerjee, P. (2016). Microbial diversity of source and point-of-use water in rural Haiti - A pyrosequencing-based metagenomic survey. *PLoS ONE*, *11*(12), e0167353. https://doi.org/10.1371/journal.pone.0167353

Navarro, S., Vela, N., & Navarro, G. (2007). Review. An overview on the environmental behaviour of pesticide residues in soils. *Spanish Journal of Agricultural Research*, (3), 357–375. https://doi.org/10.5424/sjar/2007053-5344

Neilson, A. H., & Allard, A.-S. (2012). *Organic Chemicals in the Environment: Mechanisms of Degradation and Transformation* (2nd ed.). Boca Raton, Florida: CRC Press. https://doi.org/10.1201/b12492

NFU. (2019). Decision to ban metaldehyde quashed. https://www.nfuonline.com/news/latest-news/decision-to-ban-metaldehyde-quashed/. Accessed 6 June 2020

Nicolaisen, M. H., Bælum, J., Jacobsen, C. S., & Sørensen, J. (2008). Transcription dynamics of the functional tfdA gene during MCPA herbicide degradation by Cupriavidus necator AEO106 (pRO101) in agricultural soil. *Environmental Microbiology*, *10*(3), 571–579. https://doi.org/10.1111/j.1462-2920.2007.01476.x

Nielsen, M. S., Bælum, J., Jensen, M. B., & Jacobsen, C. S. (2011). Mineralization of the herbicide MCPA in urban soils is linked to presence and growth of class III tfdA genes. *Soil Biology and Biochemistry*, *43*(5), 984–990. https://doi.org/10.1016/j.soilbio.2011.01.014

Nojiri, H., Shintani, M., & Omori, T. (2004). Divergence of mobile genetic elements involved in the distribution of xenobiotic-catabolic capacity. *Applied Microbiology and Biotechnology*, *64*(2), 154–174. https://doi.org/10.1007/s00253-003-1509-y

Notomi, T., Okayama, H., Masubuchi, H., Yonekawa, T., Watanabe, K., Amino, N., & Hase, T. (2000). Loop-mediated isothermal amplification. *Nucleic Acids Research*, *28*(12), e63. https://doi.org/10.1093/nar/28.12.e63

Nousiainen, A. O., Björklöf, K., Sagarkar, S., Mukherjee, S., Purohit, H. J., Kapley, A., & Jørgensen, K. S. (2014). Atrazine degradation in boreal nonagricultural subsoil and tropical agricultural soil. *Journal of Soils and Sediments*, *14*(6), 1179–1188. https://doi.org/10.1007/s11368-014-0868-6

Nzila, A. (2013). Update on the cometabolism of organic pollutants by bacteria. *Environmental pollution*, *178*, 474–482. https://doi.org/10.1016/j.envpol.2013.03.042

O'Brien, P. J., & Herschlag, D. (1999). Catalytic promiscuity and the evolution of new enzymatic activities. *Chemistry and Biology*, *6*(4), R91–R105. https://doi.org/10.1016/S1074-5521(99)80033-7

Ochman, H., Lawrence, J. G., & Grolsman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature*, *405*(6784), 299–304. https://doi.org/10.1038/35012500

Okonechnikov, K., Golosova, O., Fursov, M., & UGENE team. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, *28*(8), 1166–7. https://doi.org/10.1093/bioinformatics/bts091

Ortiz-Hernández, M., Sánchez-Salinas, E., Olvera-Velona, A., & Luis, J. (2011). Pesticides in the Environment: Impacts and its Biodegradation as a Strategy for Residues Treatment. In M. Stoytcheva (Ed.), *Pesticides - Formulations, Effects, Fate* (pp. 551–574). Rijeka, Croatia: InTech. https://doi.org/10.5772/13534

Oxford Nanopore Technologies. (2020). R10.3: the newest nanopore for high accuracy nanopore sequencing – now available in store. https://nanoporetech.com/about-us/news/r103-newest-nanopore-high-accuracy-nanopore-sequencing-now-available-store. Accessed 16 March 2020

Parekh, N. R., Hartmann, A., Charnay, M. P., & Fournier, J. C. (1995). Diversity of carbofuran-degrading soil bacteria and detection of plasmid-encoded sequences homologous to the mcd gene. *FEMS Microbiology Ecology*, *17*(3), 149–160. https://doi.org/10.1016/0168-6496(95)00019-7

Perruchon, C., Patsioura, V., Vasileiadis, S., & Karpouzas, D. G. (2016). Isolation and characterisation of a *Sphingomonas* strain able to degrade the fungicide *ortho* -phenylphenol. *Pest Management Science*, *72*(1), 113–124. https://doi.org/10.1002/ps.3970

Pesce, S., Beguet, J., Rouard, N., Devers-Lamrani, M., & Martin-Laurent, F. (2013). Response of a diuron-degrading community to diuron exposure assessed by real-time quantitative PCR monitoring of phenylurea hydrolase A and B encoding genes. *Applied Microbiology and Biotechnology*, *97*(4), 1661–1668. https://doi.org/10.1007/s00253-012-4318-3

Pfannes, K. R., Langenbach, K. M. W., Pilloni, G., Stührmann, T., Euringer, K., Lueders, T., et al. (2015). Selective elimination of bacterial faecal indicators in the Schmutzdecke of slow sand filtration columns. *Applied Microbiology and Biotechnology*, *99*(23), 10323–10332. https://doi.org/10.1007/s00253-015-6882-9

Piatak, M., Luk, K. C., Williams, B., & Lifson, J. D. (1993). Quantitative competitive polymerase chain reaction for accurate quantitation of HIV DNA and RNA species. *BioTechniques*, *14*(1), 70–81.

Picard, C., Ponsonnet, C., Paget, E., Nesme, X., & Simonet, P. (1992). Detection and enumeration of bacteria in soil by direct DNA extraction and polymerase chain reaction. *Applied and Environmental Microbiology*, *58*(9), 2717–2722.

Pignatello, J. J., & Xing, B. (1996). Mechanisms of slow sorption of organic chemicals to natural particles. *Environmental Science and Technology*, *30*(1), 1–11. https://doi.org/10.1021/es940683g

Piwoni, M. D., & Keeley, J. W. (1990). *Basic concepts of contaminant sorption at hazardous waste*

sites. *Ground Water Issue -EPA/540/4-90/053*. https://doi.org/10.1056/NEJMoa030660

Pizzetti, I., Fuchs, B., Gerdts, G., Wichels, A., Wiltshire, K., & Amann, R. (2011). Temporal variability of coastal Planctomycetes clades at station Kabeltonne, North Sea. *Appl. Environ. Microbiol.*, *77*(14), 5009–5017. https://doi.org/10.1128/AEM.02931-10,hdl:10013/epic.37854

Pompei, C. M. E., Ciric, L., Canales, M., Karu, K., Vieira, E. M., & Campos, L. C. (2017). Influence of PPCPs on the performance of intermittently operated slow sand filters for household water purification. *Science of the Total Environment*, *581*, 174–185. https://doi.org/10.1016/j.scitotenv.2016.12.091

Post, V., & Hall, R. M. (2009). AbaR5, a large multiple-antibiotic resistance region found in Acinetobacter baumannii. *Antimicrobial Agents and Chemotherapy*, *53*(6), 2667–2671. https://doi.org/10.1128/AAC.01407-08

PPDB. (2020). Pesticide Properties DataBase, University of Hertfordshire. https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/446.htm

Prasse, C., Wagner, M., Schulz, R., & Ternes, T. A. (2011). Biotransformation of the antiviral drugs acyclovir and penciclovir in activated sludge treatment. *Environmental Science and Technology*, *45*(7), 2761–2769. https://doi.org/10.1021/es103732y

Providenti, M. A., O'Brien, J. M., Ewing, R. J., Paterson, E. S., & Smith, M. L. (2006). The copy-number of plasmids and other genetic elements can be determined by SYBR-Green-based quantitative real-time PCR. *Journal of Microbiological Methods*, *65*(3), 476–487. https://doi.org/10.1016/j.mimet.2005.09.007

Rački, N., Dreo, T., Gutierrez-Aguirre, I., Blejec, A., & Ravnikar, M. (2014). Reverse transcriptase droplet digital PCR shows high resilience to PCR inhibitors from plant, soil and water samples. *Plant Methods*, *10*(1), 42. https://doi.org/10.1186/s13007-014-0042-6

Ramond, J. B., Welz, P. J., Tuffin, M. I., Burton, S. G., & Cowan, D. A. (2013). Assessment of temporal and spatial evolution of bacterial communities in a biological sand filter mesocosm treating winery wastewater. *Journal of Applied Microbiology*, *115*(1), 91–101. https://doi.org/10.1111/jam.12203

Ramsay, J. P., & Firth, N. (2017). Diverse mobilization strategies facilitate transfer of non-conjugative mobile genetic elements. *Current Opinion in Microbiology*, *38*, 1–9. https://doi.org/10.1016/j.mib.2017.03.003

Ranjan, P., & Prem, M. (2018). Schmutzdecke- A Filtration Layer of Slow Sand Filter. *International Journal of Current Microbiology and Applied Sciences*, *7*(7), 637–645. https://doi.org/10.20546/ijcmas.2018.707.077

Rasko, D. A., Myers, G. S., & Ravel, J. (2005). Visualization of comparative genomic analyses by BLAST score ratio. *BMC Bioinformatics*, *6*(1), 2. https://doi.org/10.1186/1471-2105-6-2

Reece, R. L., Scott, P. C., Forsyth, W. M., Gould, J. A., & Barr, D. A. (1985). Toxicity episodes involving agricultural chemicals and other substances in birds in Victoria, Australia. *Veterinary Record*, *117*(20), 525–527. https://doi.org/10.1136/vr.117.20.525

Regar, R. K., Gaur, V. K., Bajaj, A., Tambat, S., & Manickam, N. (2019). Comparative microbiome analysis of two different long-term pesticide contaminated soils revealed the anthropogenic influence on functional potential of microbial communities. *Science of the Total Environment*, *681*, 413–423. https://doi.org/10.1016/j.scitotenv.2019.05.090

Riveros-Rosas, H., González-Segura, L., Julián-Sánchez, A., Díaz-Sánchez, Á. G., & Muñoz-Clares, R. A. (2013). Structural determinants of substrate specificity in aldehyde dehydrogenases. *Chemico-Biological Interactions*, *202*(1–3), 51–61. https://doi.org/10.1016/j.cbi.2012.11.015

Riveros-Rosas, H., Julián-Sánchez, A., Moreno-Hagelsieb, G., & Muñoz-Clares, R. A. (2019). Aldehyde dehydrogenase diversity in bacteria of the Pseudomonas genus. *Chemico-Biological Interactions*, *304*, 83–87. https://doi.org/10.1016/j.cbi.2019.03.006

Roberts, S. J., Walker, A., Parekh, N. R., Welch, S. J., & Waddington, M. J. (1993). Studies on a mixed bacterial culture from soil which degrades the herbicide linuron. *Pesticide Science*, *39*(1), 71–78. https://doi.org/10.1002/ps.2780390111

Rodríguez-Cruz, M. S., Bælum, J., Shaw, L. J., Sørensen, S. R., Shi, S., Aspray, T., et al. (2010). Biodegradation of the herbicide mecoprop-p with soil depth and its relationship with class III tfdA genes. *Soil Biology and Biochemistry*, *42*(1), 32–39. https://doi.org/10.1016/j.soilbio.2009.09.018

Rolph, C. (2016). *21st Century biological processes for metaldehyde removal in drinking water treatment*. Cranfield University.

Rolph, C., Villa, R., Jefferson, B., Brookes, A., Choya, A., Iceton, G., & Hassard, F. (2019). From full-scale biofilters to bioreactors: Engineering biological metaldehyde removal. *Science of the Total Environment*, *685*, 410–418. https://doi.org/10.1016/j.scitotenv.2019.05.304

Rooklidge, S. J., Burns, E. R., & Bolte, J. P. (2005). Modeling antimicrobial contaminant removal in slow sand filtration. *Water Research*, *39*(2–3), 331–339. https://doi.org/10.1016/j.watres.2004.09.024

Rousidou, C., Karaiskos, D., Myti, D., Karanasios, E., Karas, P. A., Tourna, M., et al. (2017). Distribution and function of carbamate hydrolase genes cehA and mcd in soils: the distinct role of soil pH. *FEMS Microbiology Ecology*, *93*(1), 1–12. https://doi.org/10.1093/femsec/fiw219

Sagarkar, S., Mukherjee, S., Nousiainen, A., Björklöf, K., Purohit, H. J., Jorgensen, K. S., & Kapley, A. (2013). Monitoring bioremediation of atrazine in soil microcosms using molecular tools. *Environmental Pollution*, *172*, 108–115. https://doi.org/10.1016/j.envpol.2012.07.048

Salminen, J. M., Tuomi, P. M., & Jørgensen, K. S. (2008). Functional gene abundances (nahAc, alkB, xylE) in the assessment of the efficacy of bioremediation. *Applied Biochemistry and Biotechnology*, *151*(2–3), 638–652. https://doi.org/10.1007/s12010-008-8275-3

Samuelsen, E. D., Badawi, N., Nybroe, O., Sørensen, S. R., & Aamand, J. (2017). Adhesion to sand and ability to mineralise low pesticide concentrations are required for efficient bioaugmentation of flow-through sand filters. *Applied Microbiology and Biotechnology*, *101*(1), 411–421. https://doi.org/10.1007/s00253-016-7909-6

Sangwan, N., Lata, P., Dwivedi, V., Singh, A., Niharika, N., Kaur, J., et al. (2012). Comparative Metagenomic Analysis of Soil Microbial Communities across Three Hexachlorocyclohexane Contamination Levels. *PLoS ONE*, *7*(9), 1–12. https://doi.org/10.1371/journal.pone.0046219

Sangwan, N., Verma, H., Kumar, R., Negi, V., Lax, S., Khurana, P., et al. (2014). Reconstructing an ancestral genotype of two hexachlorocyclohexane-degrading Sphingobium species using metagenomic sequence data. *ISME Journal*, *8*(2), 398–408. https://doi.org/10.1038/ismej.2013.153

Scheytt, T., Mersmann, P., Leidig, M., Pekdeger, A., & Heberer, T. (2004). Transport of Pharmaceutically Active Compounds in Saturated Laboratory Columns. *Ground Water*, *42*(5), 767–773. https://doi.org/10.1111/j.1745-6584.2004.tb02730.x

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Seffernick, J. L., De Souza, M. L., Sadowsky, M. J., & Wackett, L. P. (2001). Melamine deaminase and atrazine chlorohydrolase: 98 percent identical but functionally different. *Journal of Bacteriology*, *183*(8), 2405–2410. https://doi.org/10.1128/JB.183.8.2405-2410.2001

Seffernick, J. L., & Wackett, L. P. (2001). Rapid evolution of bacterial catabolic enzymes: A case study with atrazine chlorohydrolase. *Biochemistry*, *40*(43), 12747–12753. https://doi.org/10.1021/bi011293r

Senesi, N. (1992). Binding mechanisms of pesticides to soil humic substances. *Science of the Total Environment*, *123*, 63–76. https://doi.org/10.1016/0048-9697(92)90133-D

Severn Trent Water. (2014). *Final Water Resources Management Plan 2014*. Coventry.

Shahi, A., Aydin, S., Ince, B., & Ince, O. (2016). Evaluation of microbial population and functional genes during the bioremediation of petroleum-contaminated soil as an effective monitoring approach. *Ecotoxicology and Environmental Safety*, *125*, 153–160. https://doi.org/10.1016/j.ecoenv.2015.11.029

Shahsavari, E., Aburto-Medina, A., Taha, M., & Ball, A. S. (2016). A quantitative PCR approach for quantification of functional genes involved in the degradation of polycyclic aromatic hydrocarbons in contaminated soils. *MethodsX*, *3*, 205–211. https://doi.org/10.1016/j.mex.2016.02.005

Shelton, D. R., & Somich, C. J. (1988). Isolation and characterization of coumaphos-metabolizing bacteria from cattle dip. *Applied and Environmental Microbiology*, *54*(10), 2566–2571. https://doi.org/10.1128/aem.54.10.2566-2571.1988

Shennan, C., Muramoto, J., Koike, S., Baird, G., Fennimore, S., Samtani, J., et al. (2018). Anaerobic soil disinfestation is an alternative to soil fumigation for control of some soilborne pathogens in strawberry production. *Plant Pathology*, *67*(1), 51–66. https://doi.org/10.1111/ppa.12721

Shintani, M., & Nojiri, H. (2013). Mobile Genetic Elements (MGEs) carrying catabolic genes. In A. Malik, E. Grohmann, & M. Alves (Eds.), *Management of Microbial Resources in the Environment* (pp. 167–214). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-5931-2_8

Siguier, P., Gourbeyre, E., & Chandler, M. (2014). Bacterial insertion sequences: Their genomic impact and diversity. *FEMS Microbiology Reviews*, *38*(5), 865–891. https://doi.org/10.1111/1574-6976.12067

Siguier, P., Gourbeyre, E., Varani, A., Ton-Hoang, B., & Chandler, M. (2015). Everyman's Guide to Bacterial Insertion Sequences. *Microbiology Spectrum*, *3*(2). https://doi.org/10.1128/microbiolspec.mdna3-0030-2014

Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., & Chandler, M. (2006). ISfinder: the reference centre for bacterial insertion sequences. *Nucleic acids research*, *34*, D32–D36. https://doi.org/10.1093/nar/gkj014

Simms, L. C., Dawson, J. J. C., Paton, G. I., & Wilson, M. J. (2006). Identification of environmental

factors limiting plant uptake of metaldehyde seed treatments under field conditions. *Journal of Agricultural and Food Chemistry*, *54*(10), 3646–3650. https://doi.org/10.1021/jf060231a

Singh, B., Kuhad, R. C., Singh, A., Lal, R., & Tripathi, K. K. (1999). Biochemical and molecular basis of pesticide degradation by microorganisms. *Critical Reviews in Biotechnology*, *19*(3), 197–225. https://doi.org/10.1080/0738-859991229242

Singh, B., Walker, A., Morgan, J. A. W., & Wright, D. J. (2004). Biodegradation of chlorpyrifos by Enterobacter strain B-14 and its use in bioremediation of contaminated soils. *Applied and Environmental Microbiology*, *70*(8), 4855–4863. https://doi.org/10.1128/AEM.70.8.4855-4863.2004

Singh, D., Prabha, R., Gupta, V. K., & Verma, M. K. (2018). Metatranscriptome analysis deciphers multifunctional genes and enzymes linked with the degradation of aromatic compounds and pesticides in the wheat rhizosphere. *Frontiers in Microbiology*, *9*, 1–15. https://doi.org/10.3389/fmicb.2018.01331

Smets, W., Leff, J. W., Bradford, M. A., McCulley, R. L., Lebeer, S., & Fierer, N. (2016). A method for simultaneous measurement of soil bacterial abundances and community composition via 16S rRNA gene sequencing. *Soil Biology and Biochemistry*, *96*, 145–151. https://doi.org/10.1016/J.SOILBIO.2016.02.003

Smillie, C., Garcillan-Barcia, M., Francia, M. V., Rocha, E. P. C., & de la Cruz, F. (2010). Mobility of Plasmids. *Microbiology and Molecular Biology Reviews*, *74*(3), 434–452. https://doi.org/10.1128/mmbr.00020-10

Sniegowski, K., Bers, K., Van Goetem, K., Ryckeboer, J., Jaeken, P., Spanoghe, P., & Springael, D. (2011). Improvement of pesticide mineralization in on-farm biopurification systems by bioaugmentation with pesticide-primed soil. *FEMS Microbiology Ecology*, *76*(1), 64–73. https://doi.org/10.1111/j.1574-6941.2010.01031.x

Solovyev, V., Salamov, A., Seledtsov, I., Vorobyev, D., & Bachinsky, A. (2011). Automatic annotation of bacterial community sequences and application to infections diagnostic. In *BIOINFORMATICS 2011 - Proceedings of the International Conference on Bioinformatics Models, Methods and Algorithms* (pp. 346–353). https://doi.org/10.5220/0003333703460353

Sørensen, S. R., Simonsen, A., & Aamand, J. (2009). Constitutive mineralization of low concentrations of the herbicide linuron by a Variovorax sp. strain. *FEMS Microbiology Letters*, *292*(2), 291–296. https://doi.org/10.1111/j.1574-6968.2009.01501.x

Springael, D., Van Thor, J., Goorissen, H., Ryngaert, A., De Baere, R., Van Hauwe, P., et al. (1996). RP4::Mu3A-mediated in vivo cloning and transfer of a chlorobiphenyl catabolic pathway. *Microbiology*, *142*(11), 3283–3293. https://doi.org/10.1099/13500872-142-11-3283

Stoddard, S. F., Smith, B. J., Hein, R., Roller, B. R. K., & Schmidt, T. M. (2015). rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Research*, *43*(Database issue), D593-8. https://doi.org/10.1093/nar/gku1201

Streminska, M. A., van der Wurff, A. W. G., Runia, W. T., Thoden, T. C., Termorshuizen, A. J., & Feil, H. (2014). Changes in bacterial and fungal abundance in soil during the process of anaerobic soil disinfestation. *Acta Horticulturae*, (1041), 95–102. https://doi.org/10.17660/ActaHortic.2014.1041.9

Stuart, M. E., Talbot, J. C., Manamsa, K., & Crane E J. (2011). *Emerging contaminants in groundwater*. Nottingham.

Tamura, K., Nei, M., & Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(30), 11030–11035. https://doi.org/10.1073/pnas.0404206101

Tao, B., & Fletcher, A. J. (2013). Metaldehyde removal from aqueous solution by adsorption and ion exchange mechanisms onto activated carbon and polymeric sorbents. *Journal of Hazardous Materials*, *244–245*, 240–250. https://doi.org/10.1016/J.JHAZMAT.2012.11.014

Tavakoli, N., Comanducci, A., Dodd, H. M., Lett, M. C., Albiger, B., & Bennett, P. (2000). IS1294, a DNA element that transposes by RC transposition. *Plasmid*, *44*(1), 66–84. https://doi.org/10.1006/plas.1999.1460

Taylor, S. C., Carbonneau, J., Shelton, D. N., & Boivin, G. (2015). Optimization of Droplet Digital PCR from RNA and DNA extracts with direct comparison to RT-qPCR: Clinical implications for quantification of Oseltamivir-resistant subpopulations. *Journal of Virological Methods*, *224*, 58–66. https://doi.org/10.1016/j.jviromet.2015.08.014

Taylor, S. C., Laperriere, G., & Germain, H. (2017). Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: From variable nonsense to publication quality data. *Scientific Reports*, *7*(1), 1–8. https://doi.org/10.1038/s41598-017-02217-x

Taylor, S. C., Nadeau, K., Abbasi, M., Lachance, C., Nguyen, M., & Fenrich, J. (2019). The Ultimate qPCR Experiment: Producing Publication Quality, Reproducible Data the First Time. *Trends in Biotechnology*, *37*(7), 761–774. https://doi.org/10.1016/j.tibtech.2018.12.002

Techtmann, S. M., & Hazen, T. C. (2016). Metagenomic applications in environmental monitoring and bioremediation. *Journal of Industrial Microbiology and Biotechnology*, *43*(10), 1345–1354. https://doi.org/10.1007/s10295-016-1809-8

Thomas, J. C. (2016). *The molecular basis for the biotic degradation of metaldehyde*. University of York.

Thomas, J. C., Helgason, T., Sinclair, C. J., & Moir, J. W. B. (2017). Isolation and characterization of metaldehyde-degrading bacteria from domestic soils. *Microbial Biotechnology*, *10*(6), 1824–1829. https://doi.org/10.1111/1751-7915.12719

Thompson, B. M., Lin, C. H., Kremer, R. J., Lerc, R. N., Hsieh, H. Y., & Garrett, H. E. (2010). Evaluation of PCR-based quantification techniques to estimate the abundance of atrazine chlorohydrolase gene atzA in rhizosphere soils. *Journal of Environmental Quality*, *39*(6), 1999–2005. https://doi.org/10.2134/jeq2010.0192

Toleman, M. A., Bennett, P. M., & Walsh, T. R. (2006). ISCR elements: novel gene-capturing systems of the 21st century? *Microbiology and Molecular Biology Reviews*, *70*(2), 296–316. https://doi.org/10.1128/MMBR.00048-05

Top, E. M., Springael, D., & Boon, N. (2002). Catabolic mobile genetic elements and their potential use in bioaugmentation of polluted soils and waters. *FEMS Microbiology Ecology*, *42*(2), 199–208. https://doi.org/10.1111/j.1574-6941.2002.tb01009.x

Treangen, T. J., & Rocha, E. P. C. (2011). Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLoS Genetics*, *7*(1), e1001284. https://doi.org/10.1371/journal.pgen.1001284

Triebskorn, R., Christensen, K., & Heim, I. (1998). Effects of orally and dermally applied metaldehyde on mucus cells of slugs (Deroceras reticulatum) depending on temperature and duration of exposure. *Journal of Molluscan Studies*, *64*(4), 467–487. https://doi.org/10.1093/mollus/64.4.467

Tuomi, P. M., Salminen, J. M., & Jørgensen, K. S. (2004). The abundance of nahAc genes correlates with the 14C-naphthalene mineralization potential in petroleum hydrocarbon-contaminated oxic soil layers. *FEMS Microbiology Ecology*, *51*(1), 99–107. https://doi.org/10.1016/j.femsec.2004.07.011

Van Der Meer, J. R., & Sentchilo, V. (2003). Genomic islands and the evolution of catabolic pathways in bacteria. *Current Opinion in Biotechnology*, *14*(3), 248–254. https://doi.org/10.1016/S0958-1669(03)00058-2

Van Eerd, L. L., Hoagland, R. E., Zablotowicz, R. M., & Hall, J. C. (2003). Pesticide metabolism in plants and microorganisms. *Weed Science*, *51*(4), 472–495. https://doi.org/10.1614/0043-1745(2003)051[0472:pmipam]2.0.co;2

Vidali, M. (2001). Bioremediation. An overview. *Pure and Applied Chemistry*, *73*(7), 1163–1172. https://doi.org/10.1351/pac200173071163

Vishniac, W., & Santer, M. (1957). The thiobacilli. *Bacteriological reviews*, *21*(3), 195–213.

Vogel, T. M. (1996). Bioaugmentation as a soil bioremediation approach. *Current Opinion in Biotechnology*, *7*(3), 311–316. https://doi.org/10.1016/S0958-1669(96)80036-X

Waite, D. T., Cessna, A. J., Grover, R., Kerr, L. A., & Snihura, A. D. (2002). Environmental Concentrations of Agricultural Herbicides. *Journal of Environmental Quality*, *31*(1), 129–144. https://doi.org/10.2134/jeq2002.1290

Walia, S., Khan, A., & Rosenthal, N. (1990). Construction and applications of DNA probes for detection of polychlorinated biphenyl-degrading genotypes in toxic organic-contaminated soil environments. *Applied and Environmental Microbiology*, *56*(1), 254–259.

Walters, W., Hyde, E. R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., et al. (2016). Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems*, *1*(1), e00009-15. https://doi.org/10.1128/mSystems.00009-15

Wang, Q., Xie, S., & Hu, R. (2013). Bioaugmentation with Arthrobacter sp. Strain DAT1 for remediation of heavily atrazine-contaminated soil. *International Biodeterioration and Biodegradation*, *77*, 63–67. https://doi.org/10.1016/j.ibiod.2012.11.003

Wattam, A. R., Davis, J. J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., et al. (2017). Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Research*, *45*(D1), D535–D542. https://doi.org/10.1093/nar/gkw1017

Webster, T. M., & Fierer, N. (2019). Microbial dynamics of biosand filters and contributions of the microbial food web to effective treatment of wastewater-impacted water sources. *Applied and Environmental Microbiology*, *85*(17), e01142-19. https://doi.org/10.1128/AEM.01142-19

Wick, R. R., Judd, L. M., Gorrie, C. L., & Holt, K. E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Computational Biology*, *13*(6), e1005595. https://doi.org/10.1371/journal.pcbi.1005595

Widada, J., Nojiri, H., & Omori, T. (2002). Recent developments in molecular techniques for

identification and monitoring of xenobiotic-degrading bacteria and their catabolic genes in bioremediation. *Applied Microbiology and Biotechnology*, *60*(1–2), 45–59. https://doi.org/10.1007/s00253-002-1072-y

Wyndham, R. C., Cashore, A. E., Nakatsu, C. H., & Peel, M. C. (1994). Catabolic transposons. *Biodegradation*, *5*(3–4), 323–342. https://doi.org/10.1007/BF00696468

Yale, R. L., Sapp, M., Sinclair, C. J., & Moir, J. W. B. (2017). Microbial changes linked to the accelerated degradation of the herbicide atrazine in a range of temperate soils. *Environmental Science and Pollution Research*, *24*(8), 7359–7374. https://doi.org/10.1007/s11356-017-8377-y

Yan, X., Gu, T., Yi, Z., Huang, J., Liu, X., Zhang, J., et al. (2016). Comparative genomic analysis of isoproturon-mineralizing sphingomonads reveals the isoproturon catabolic mechanism. *Environmental Microbiology*, *18*(12), 4888–4906. https://doi.org/10.1111/1462-2920.13413

Yang, G., Anderson, D. W., Baier, F., Dohmen, E., Hong, N., Carr, P. D., et al. (2019). Higher-order epistasis shapes the fitness landscape of a xenobiotic-degrading enzyme. *Nature Chemical Biology*, *15*(11), 1120–1128. https://doi.org/10.1038/s41589-019-0386-3

Yang, L., Li, X., Li, X., Su, Z., Zhang, C., & Zhang, H. (2014). Bioremediation of chlorimuron-ethyl-contaminated soil by Hansschlegelia sp. strain CHL1 and the changes of indigenous microbial population and N-cycling function genes during the bioremediation process. *Journal of Hazardous Materials*, *274*, 314–321. https://doi.org/10.1016/j.jhazmat.2014.04.011

Zaprasis, A., Liu, Y. J., Liu, S. J., Drake, H. L., & Horn, M. A. (2010). Abundance of novel and diverse tfdA-like genes, encoding putative phenoxyalkanoic acid herbicide-degrading dioxygenases, in soil. *Applied and Environmental Microbiology*, *76*(1), 119–128. https://doi.org/10.1128/AEM.01727-09

Zearley, T. L., & Summers, R. S. (2012). Removal of trace organic micropollutants by drinking water biological filters. *Environmental Science and Technology*, *46*(17), 9412–9419. https://doi.org/10.1021/es301428e

Zhalnina, K., de Quadros, P. D., Gano, K. A., Davis-Richardson, A., Fagen, J. R., Brown, C. T., et al. (2013). Ca. Nitrososphaera and Bradyrhizobium are inversely correlated and related to agricultural practices in long-term field experiments. *Frontiers in Microbiology*, *4*, 104. https://doi.org/10.3389/fmicb.2013.00104

Zhang, H., Wang, C., Lu, H., Guan, W., & Ma, Y. (2011). Residues and dissipation dynamics of molluscicide metaldehyde in cabbage and soil. *Ecotoxicology and Environmental Safety*, *74*(6), 1653–1658. https://doi.org/10.1016/J.ECOENV.2011.05.004

Zhang, X., & Dai, X. (2006). Degradation and determination of the residue of metaldehyde in tobacco and soil. *Chinese Journal of Pesticide Science*, *4*, 344–348.

Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000). A Greedy Algorithm for Aligning DNA Sequences. *Journal of Computational Biology*, *7*(1–2), 203–214. https://doi.org/10.1089/10665270050081478

Zimmermann, K., & Mannhalter, J. W. (1996). Technical aspects of quantitative competitive PCR. *BioTechniques*, *21*(2), 268–279. https://doi.org/10.2144/96212rv01

Zuehlke, S., Duennbier, U., & Heberer, T. (2007). Investigation of the behavior and metabolism of pharmaceutical residues during purification of contaminated ground water used for

drinking water supply. *Chemosphere*, *69*(11), 1673–1680. https://doi.org/10.1016/j.chemosphere.2007.06.020

Annex 1: Published Manuscript Soil Biology and Biochemistry 142 (2020) 107702 doi:10.1016/j.soilbio.2019.107702

Soil Biology and Biochemistry 142 (2020) 107702

Contents lists available at ScienceDirect

Soil Biology and Biochemistry

journal homepage: http://www.elsevier.com/locate/soilbio

Genomic basis for pesticide degradation revealed by selection, isolation and characterisation of a library of metaldehyde-degrading strains from soil

Víctor Castro-Gutiérrez [a,b], Edward Fuller [a], John C. Thomas [a,1], Chris J. Sinclair [c], Steven Johnson [d], Thorunn Helgason [a], James W.B. Moir [a,*]

[a] Department of Biology, University of York, Heslington, York, UK
[b] Centro de Investigación en Contaminación Ambiental (CICA), Universidad de Costa Rica, San José, Costa Rica
[c] FERA Science Ltd (Fera), National Agri-Food Innovation Campus, Sand Hutton, York, UK
[d] Department of Electronic Engineering, University of York, Heslington, York, UK

ABSTRACT

Metaldehyde, a xenobiotic cyclic ether, is used as molluscicide of choice in agriculture and horticulture, but recently its detection in drinking water sources has become a major cause of concern. We isolated eight new metaldehyde-degrading bacterial strains from allotment and agricultural soils and identified a highly-conserved gene cluster shared amongst one gamma and five beta-proteobacteria, and absent from closely-related, non-degrading type strains. Chemical mutagenesis, and heterologous expression in *E. coli*, confirmed that this gene cluster is responsible for metaldehyde degradation. Other metaldehyde-degrading isolates that lack this pathway indicate that multiple degradation mechanisms have evolved. We demonstrated accelerated biodegradation of metaldehyde in multiple soils, highlighting the importance of the biological component in metaldehyde degradation in nature. We confirmed that the metaldehyde-degrading population in soil is proliferating in response to metaldehyde, but no bulk changes in the composition of the community as a whole were detected, indicating the process is governed by a few rare taxa. Here, we identified the first genetic determinants for the biological degradation of metaldehyde in soil paving the way for targeted bioremediation strategies.

1. Introduction

Soils contain extremely diverse microbial communities with versatile metabolic capabilities. The acquisition of new traits occurs by diversification of the existing genetic material from the metagenome and is further enabled by horizontal gene transfer (Maheshwari et al., 2017). Thus, it is possible for novel metabolic activities to emerge and be evolutionarily reinforced via selection within soil microbial communities (Fierer, 2017; Kuzyakov and Blagodatskaya, 2015). Such processes are important to agriculture as pesticides are used to improve crop productivity, but also impact soil microbes, which may degrade them. Genes whose products degrade novel chemicals and confer a competitive nutritional advantage on their microbial hosts have potential to be selected for in soils (Arbeli and Fuentes, 2007).

Metaldehyde ($CH_3CHO)_4$ is a pesticide used worldwide as a molluscicide to control snails and slugs that damage agricultural crops and domestic gardens (Eckert et al., 2012). It is a xenobiotic that has been used as a pesticide since the 1930s (Gimingham, 1940); it is hydrolytically and photolytically stable (Carpenter, 1989; Kegley et al., 2016). Metaldehyde is applied to many crops, including oilseed rape, wheat and potatoes, and accounted for 84.5% of all molluscicide applications in the UK in 2016 (Garthwaite et al., 2018). It is applied as a pelleted bran bait in the autumn, when molluscs prosper due to the humid conditions. Rainfall can dissolve metaldehyde and carry the compound to watercourses which may be used for drinking water abstraction (Lazartigues et al., 2013).

Metaldehyde is recalcitrant to removal using conventional drinking water treatment processes based on adsorption of substances to activated carbon (Castle et al., 2017). Moreover, detections of metaldehyde have been the main cause for already treated water not meeting pesticide standards since its monitoring began (Chief Inspector of Drinking Water, 2017). Hence, water companies have been exploring alternative solutions for metaldehyde detection and elimination. Therefore, the idea of using biological strategies to detect and degrade metaldehyde has recently emerged. Metaldehyde can be quickly degraded in soils (Lewis et al., 2016; Zhang et al., 2011) and is oxidised to carbon dioxide under aerobic conditions (European Food Safety Authority, 2010), in comparison to its long half-life in sterile conditions (Simms et al., 2006). This

* Corresponding author.
E-mail address: james.moir@york.ac.uk (J.W.B. Moir).
[1] Current address: Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge, UK.

https://doi.org/10.1016/j.soilbio.2019.107702
Received 25 September 2019; Received in revised form 22 December 2019; Accepted 29 December 2019
Available online 2 January 2020

suggests a strong involvement of microorganisms in its degradation. The biotic degradation of metaldehyde in treated soil or sediments proceeds at very different rates in different experiments; with DT_{50} values of between <1 and 67 d (Cranor, 1990; Möllerfeld et al., 1993; Zhang and Dai, 2006).

Very few studies have investigated further the mechanism underpinning the microbial degradation of metaldehyde. Rolph (2016) found improved metaldehyde removal in slow sand filter material following a period of acclimation with relatively elevated metaldehyde concentrations. Thomas and collaborators (Thomas et al., 2017) were the first to isolate bacterial strains capable of metaldehyde degradation (*Acinetobacter calcoaceticus* E1 and *Variovorax* E3). Even so, the diversity of metaldehyde degraders found in nature, the biochemical mechanisms responsible for its degradation, and the effect metaldehyde inputs have in shaping microbial communities in biological habitats remain largely unexplored. Here we examine the diversity of organisms capable of degrading metaldehyde using culture-based and culture-independent molecular methods, focusing on the impact of historical and experimental application of the pesticide. We isolate a set of genes, shared amongst several independent isolates, responsible for metaldehyde degradation.

2. Materials and methods

2.1. Chemicals and reagents

Metaldehyde (99%) was purchased from Acros Organics, NJ; sodium sulfate (≥99%) was purchased from Honeywell-Fluka, Bucharest; all other chemicals were purchased from Sigma-Aldrich, St. Louis, MO.

2.2. Analytical methods for metaldehyde

2.2.1. Metaldehyde extraction and quantification in aqueous matrices

Aqueous samples were centrifuged (4000 g, 10 min) and 0.4 mL supernatant added to 0.5 mL dichloromethane in glass chromatography vials (Thermo Scientific), vortexed for 30 s and stabilized for 20 min. 5 μL of organic phase were injected into an Agilent 7820A gas chromatograph (Stockport, UK) fitted with a HP-5 column and Flame Ionization Detector. Chromatographic parameters were described previously (Tao and Fletcher, 2013). Limit of detection and limit of quantification were 0.05 mg L^{-1} and 0.15 mg L^{-1} respectively. Calibration curves were constructed with standard metaldehyde solutions in minimal salts media (MSM) (Thomas et al., 2017) and sample peak area was interpolated in the calibration curves.

2.2.2. Metaldehyde extraction and quantification in soil

Each sample for extraction consisted of 10 g of soil in glass Falcon tubes (Kimble-Chase, NJ). 5 g anhydrous sodium sulfate were added, contents homogenized, and 15 mL dichloromethane added. Tubes were vortexed for 30 s and placed in ultrasonic bath for 20 min. Centrifugation (2000 rpm, 10 min) followed. 2 mL of supernatant were withdrawn from each system, passed through sodium sulfate/silanized glass wool column and filtered through a 0.45 μm PTFE filter. Chromatographic analysis was as described for aqueous samples. Peak area for samples was interpolated in a calibration curve. Limit of detection and limit of quantification were 0.25 mg kg^{-1} soil and 0.75 mg kg^{-1} soil respectively.

2.3. Soil sampling and characterization

Three soil sample collections were performed. First collection, January 2017: two council-run allotment plots (Hob Moor Allotments, York, UK; 53.946113, −1.103654). Plot J-Met had been regularly treated with metaldehyde for at least 5 years. Plot J-NoMet had not been treated with metaldehyde for at least 5 years. Second collection, June 2017: a lettuce and cilantro-producing agricultural farm (Caballo Blanco, Cartago, Costa Rica; 9.854884, −83.901134) periodically exposed to metaldehyde for at least 3 years (C-Met). Third collection, October 2017: two different allotment sites: Scarcroft Allotments (S) (York, UK; 53.950840, −1.092036) and Hob Moor Allotments (H). For each site, two strawberry plots were chosen: one treated with metaldehyde during the July–August period (S-Met and H-Met) and another plot with no metaldehyde applications for at least 5 years (S-NoMet, H-NoMet). In each plot, triplicate 300 g sub-samples were taken from the top 10 cm of the soil and stored in plastic bags. These were combined into a composite sample and stored in loosely tied plastic bags in a temperature-controlled room at 23 °C until analysed. Soil samples were air dried overnight and sieved (2 mm mesh). Physical and chemical parameters for soils are shown in Table S1.

2.4. Microbiological analyses

Fig. 1 shows an overview of methods used for the microbiological analyses of the collected soil samples. Soils J-Met, J-NoMet and C-Met were used for isolation of metaldehyde-degrading bacterial strains only. Soils S-Met, S-NoMet, H-Met, H-NoMet were additionally used for constructing soil metaldehyde degradation profiles and analysing associated changes in the microbial communities.

2.4.1. Selective enrichment for the isolation of metaldehyde-degrading strains (a)

A modified selective-enrichment procedure (Abraham and Silambarasan, 2013) was performed to obtain metaldehyde-degrading bacterial strains from the soils. For the first enrichment passage (P1), 1 g of soil was inoculated in 100 mL of MSM with added metaldehyde as sole source of carbon at 100 mg L^{-1} in 250 mL Erlenmeyer flasks. All metaldehyde-containing media was filter sterilized (0.22 μm) prior to use and supplemented with 2.0 mL L^{-1} trace elements (Vishniac and Santer, 1957). P1 was incubated with orbital shaking (25 °C, 150 rpm) in the dark for 72 h. Subsequently, 1.0 mL was transferred to 100 mL fresh medium (P2) and incubated as described for P1. Third and fourth passages (P3 and P4) were carried out similarly. Aliquots from P3 and P4 were streaked onto solid supplemented MSM (0.75% agarose) with 75 mg L^{-1} metaldehyde in triplicate and incubated for 72 h at 25 °C in the dark. Different colony morphotypes that grew on this media were purified by restreaking. Pure cultures were also streaked on MSM plates with no added carbon source to identify possible agarose degraders and oligotrophs, which were discarded after their inability to degrade metaldehyde was corroborated in liquid media. Strains that grew on metaldehyde plates and not in no added carbon plates were presumptively identified as metaldehyde degraders. Isolates were preserved by freezing at −80 °C in 15% glycerol.

2.4.2. Metaldehyde degradation profiles and microbial community changes in allotment soil microcosms (b)

Metaldehyde degradation was followed in allotment soil samples S-Met, S-NoMet, H-Met, H-NoMet after single and repeated metaldehyde applications at 15 mg kg^{-1} soil to investigate accelerated degradation of metaldehyde. Samples were taken for soil genomic DNA extraction to determine changes in soil microbial community in response to metaldehyde addition and incubation of sieved soils. Controls with no metaldehyde addition were labelled cS-Met, cS-NoMet, cH-Met and cH-NoMet. Details are described in Supplementary Materials and Methods.

2.4.3. 16S rRNA gene amplicon sequencing analyses (c)

Genomic DNA was extracted from the original soils, from soils during metaldehyde degradation assays and from passage P4 of all selective enrichments for community 16S rRNA gene amplicon sequence analyses, as well as from the pure cultures of metaldehyde-degrading strains for identification. Details of DNA extraction, amplification and sequencing of 16S rRNA genes are described in Supplementary Materials and Methods. A relative 16S rRNA gene copy number quantification

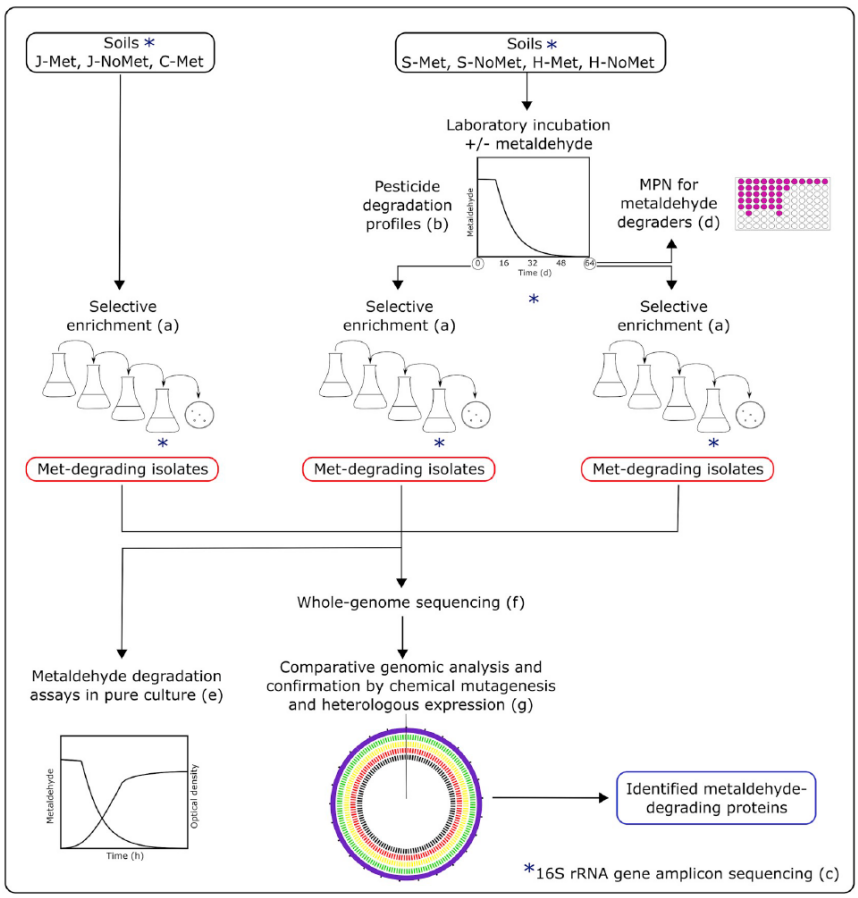


Fig. 1. Overview of microbiological analyses performed for soil samples. All soils were subjected to a selective enrichment process in minimal medium with metaldehyde as the only source of carbon for the isolation of metaldehyde-degrading strains (**a**). Specific soils were also incubated in the laboratory with and without an initial addition of metaldehyde; remaining pesticide concentrations were periodically quantified for the construction of soil metaldehyde degradation profiles (**b**). 16S rRNA gene amplicon sequencing analyses (**c**) were performed for the allotment soils before, during and after the laboratory incubations with and without metaldehyde to evaluate changes in the microbial communities as a consequence of metaldehyde exposure. After incubation, the number of metaldehyde-degrading microorganisms was determined using a most-probable number technique (**d**). Soils subjected to laboratory incubation were also used for selective enrichment to isolate metaldehyde-degrading strains. Metaldehyde-degrading capabilities of the isolated strains were tested through metaldehyde degradation assays in pure culture (**e**). Whole-genome sequencing was performed for the metaldehyde-degrading strains (**f**); comparative genomic analysis (**g**) was carried out for the identification of candidate genes involved in metaldehyde-degradation pathways and confirmed by chemical mutagenesis and heterologous expression assays.

technique (Smets et al., 2016) was carried for soils by spiking them before extraction with an internal DNA standard (*Thermus thermophilus* DSMZ 46338) at an estimated 1% of total DNA. Whole-community 16S rRNA gene amplicon sequence information was analysed using QIIME2 v2017.12 (Caporaso et al., 2010) as described in further detail in Supplementary Materials and Methods.

PRIMER7 (Primer-E Ltd., Auckland, New Zealand) was used for all statistical analyses. Abundance data was standardized by samples, transformed by fourth root and a Bray-Curtis similarity matrix was constructed. Principal coordinates analysis (PCO) was used for data ordination. Permutational MANOVA (PERMANOVA) was used assess the influence of different factors in community composition (9999 permutations). PERMDISP was used to test for homogeneity of multivariate dispersions.

2.4.4. *Enumeration of metaldehyde-degrading microorganisms in allotment soils (d)*

The number of aerobic culturable metaldehyde-degrading microorganisms in allotment soil samples after laboratory incubation±metaldehyde was determined using a most probable number (MPN) enumeration technique in microtitre plates (Dinamarca et al., 2007). Allotment soils S-Met, S-NoMet, H-Met and H-NoMet exposed to metaldehyde during metaldehyde degradation assays and their respective no metaldehyde controls (cS-Met, cS-NoMet, cH-Met and cH-NoMet) were analysed.

2.4.5. *Metaldehyde degradation assays in pure culture (e)*

To confirm the ability of isolated strains to degrade metaldehyde, bacterial growth and metaldehyde degradation assays were performed in MSM with 150 mg L^{-1} metaldehyde. Inocula were prepared by

growing each strain on nutrient agar for 72 h and resuspending growth in MSM to initial OD_{600nm} = 0.1. 0.625 mL of this inoculum were added to 100 mL metaldehyde-supplemented MSM in triplicate 250 mL Erlenmeyer flasks. Abiotic controls with an equal volume of MSM instead of inoculum were also prepared. Cultures were incubated in an orbital shaker at 25 °C and 150 rpm in the dark. Samples (1 mL) were withdrawn from the triplicate independent cultures at different time points for OD_{600nm} measurement and metaldehyde quantification.

2.4.6. Whole-genome sequencing of metaldehyde-degrading strains (f)

Confirmed metaldehyde-degrading strains were sent for whole-genome sequencing at MicrobesNG (Birmingham, UK). Sequencing was performed on Illumina MiSeq platform using 2x250bp paired-end reads. *De novo* assembly and quality assessment were performed using SPAdes (Bankevich et al., 2012) and QUAST (Gurevich et al., 2013) respectively. Automated fast annotation was performed using Prokka (Seemann, 2014).

2.4.7. Identification of candidate genes involved in metaldehyde-degradation pathways (g)

To identify candidates for proteins involved in the metaldehyde-degradation pathway, the annotated genomes from newly isolated metaldehyde-degrading strains (*Acinetobacter bohemicus* JMET-C, *Acinetobacter lwoffii* SMET-C, *Pseudomonas vancouverensis* SMET-B, *Caballeronia jiangsuensis* SNO-D), that of the previously identified degrader *A. calcoaceticus* strain E1 (Thomas et al., 2017) and closely related non metaldehyde-degrading reference strains (*A. calcoaceticus* RUH2202 and *A. bohemicus* ANC3994) were compared. Inability of these latter reference strains to degrade metaldehyde was corroborated experimentally. Reference strains were purchased from the Leibniz Institute DSMZ culture collection. A previously developed Python script (JC Thomas; available at https://pypi.org/project/blast-score-ratio/) was used to identify proteins shared between degrading strains but absent from the non-degrading strains. Using this tool, BLAST score ratio (BSR) (Rasko et al., 2005) was calculated for each of the annotated proteins of *A. calcoaceticus* E1 against the most similar proteins present in each of the other degrading and non-degrading *Acinetobacter* strains, *P. vancouverensis* SMET-B, and *C. jiangsuensis* SNO-D. Proteins with a BSR value equal to 0.9 or more (shared between the degrading strains) and with a BSR value lower than 0.45 in the non-degrading strains were chosen as candidate proteins and listed. The results of the proteome comparison analyses were corroborated, and figures generated using the BSR-based PATRIC Proteome Comparison Service (Wattam et al., 2017).

For confirmation, chemical mutagenesis was carried out by culturing A. *calcoaceticus* E1 in 20 mL LB media to an OD_{600nm} = 0.5, adding 180 μL of ethyl methanesulfonate (EMS) and incubating at 30 °C for 3 h without shaking. 500 μL of sample were inoculated into 40 mL of LB media and incubated at 30 °C for 4 h. Samples were then plated onto LB agar at various dilutions and incubated for 2 days at 30 °C (Geißdörfer et al., 1999). Single colonies were picked and transferred to MSM plates containing either acetate or metaldehyde. Mutants unable to grow on metaldehyde were sent for whole-genome sequencing (as above). Mutations in genes shared between *A. calcoaceticus* E1 and *A. bohemicus* JMET-C but absent from the respective type strains were identified by BLASTN.

Heterologous expression of putative metaldehyde degradation genes was carried out in *E. coli*. Genomic regions from *A. calcoaceticus* E1 containing *mahX*, *mahX* + *mahY*, and *mahX* + *mahY* + *aldH* were amplified with various sets of primers (Table S8) and inserted into the *Eco*RI site from pBR322. Plasmids were transformed into *E. coli* DH5α, cultured in liquid LB media, and metaldehyde disappearance measured using Gas Chromatography, as described above.

2.5. Accession numbers

Raw reads for 16S rRNA amplicons and draft whole-genome sequencing data for metaldehyde-degrading strains were deposited in the European Nucleotide Archive under study PRJEB30540. Specific sequences for metaldehyde-degrading genes are available as Supplementary Material.

3. Results

3.1. Initial enrichment cultures and sporadic isolation of metaldehyde-degrading strains

An overview of the experimental approach used in this work is shown in Fig. 1. A selective enrichment procedure in liquid media with metaldehyde as sole source of carbon was applied to soil samples from three separate soil collections to isolate metaldehyde-degrading strains. Changes in the abundance of taxa generated by this process were assessed by 16S rRNA gene amplicon sequence analysis (Table S2). The complete list of metaldehyde-degrading strains successfully isolated by the enrichment culture procedure throughout the whole study and their respective identification is presented in Table 1. Not surprisingly, the taxa corresponding to the subsequently isolated metaldehyde-degrading strains tended to dominate the composition of the final stage of the enrichment cultures. However, in some cases other taxa made up a considerable proportion of the community. These may constitute metaldehyde-degrading strains that cannot be isolated in solid media using this approach or non-degrading strains that are using metaldehyde degradation products or other by-products of degrading strains as sources of carbon (Neilson and Allard, 2012).

For soils from first collection, isolation of a metaldehyde-degrading strain (*A. bohemicus* strain JMET-C) was successful from metaldehyde-exposed soil (J-Met) (Table 1), while no metaldehyde-degrading strains were isolated from the non-exposed soil (J-NoMet). For the second soil collection, performed in a previously-exposed agricultural soil from Costa Rica (soil C-Met), metaldehyde-degrading strain *Sphingobium* sp. strain CMET-H was successfully isolated. For allotment soils from the third collection, which included metaldehyde exposed and non-exposed soils (H-Met, S-Met, H-NoMet, S-NoMet), no metaldehyde-degrading strains were initially isolated. Thus, at that point no degraders could be isolated from non metaldehyde-exposed soils and the isolation of metaldehyde degraders from soils previously exposed to metaldehyde was sporadic.

3.2. Metaldehyde is degraded faster in soils after metaldehyde treatment

Metaldehyde degradation profiles for allotment soils H-Met, H-NoMet, S-Met and S-NoMet after a single application are shown in Fig. 2a. To quantify the persistence of metaldehyde in the soil samples, data regression for the degradation profiles was performed using Single First Order or modified Hockey-stick models (Table S3) (FOCUS, 2006). With a half-life of 3.9 d, degradation was much faster in soil S-Met than other soils.

Fig. 2b shows metaldehyde degradation profiles for allotment soils subjected to two consecutive pesticide applications after a 6-month storage. A single pesticide application was enough to generate 1.7–6.1-fold reductions in metaldehyde half-lives in all the allotment soil samples on the second application; however, the effect was more evident in the soils in which the half-life was initially higher. The results show accelerated degradation occurred in the samples, which highlights the importance of biological mechanisms for metaldehyde elimination in soil.

3.3. Consistent enrichment and isolation of a greater diversity of metaldehyde degraders after incubating soil microcosms with metaldehyde in the laboratory

The accelerated degradation of metaldehyde pointed towards evolutionary selection of biologically-driven degradation of

Table 1
Metaldehyde-degrading strains isolated by enrichment culture procedure and identification based on 16S rRNA gene sequences against the NCBI database (limited to sequences from type material).

Soil of origin	Strain code	Closest relative	Closest relative GenBank Accession No.	Similarity (%)	No. of bases compared	Identification of the isolate
J-Met	JMET-C	*Acinetobacter bohemicus* ANC 3994(T)	KB849175	99.9	1406	*Acinetobacter bohemicus*
C-Met	CMET-H	*Sphingobium chlorophenolicum* NBRC 16172	NR_113840.1	98.5	1329	*Sphingobium* sp.
H-Met	HMET-A	*Rhodococcus globerulus* NBRC 14531(T)	BCWX01000023	100.0	1366	*Rhodococcus globerulus*
H-Met	HMET-G	*Sphingobium chlorophenolicum* NBRC 16172	NR_113840.1	98.4	1356	*Sphingobium* sp.
H-NoMet	HNO-A	*Rhodococcus globerulus* NBRC 14531(T)	BCWX01000023	100.0	1370	*Rhodococcus globerulus*
S-Met	SMET-B	*Pseudomonas vancouverensis* ATCC 700688(T)	AJ011507	98.9	1387	*Pseudomonas vancouverensis*
S-Met	SMET-C	*Acinetobacter lwoffii* NCTC 5866(T)	AIEL01000120	98.7	1394	*Acinetobacter lwoffii*
S-NoMet	SNO-D	*Burkholderia jiangsuensis* MP-1[a]	NR_133991.1	99.9	1398	*Caballeronia jiangsuensis*

[a] The taxon *Burkholderia jiangsuensis* has been reclassified as *Caballeronia jiangsuensis* (Dobritsa and Samadpour, 2016).

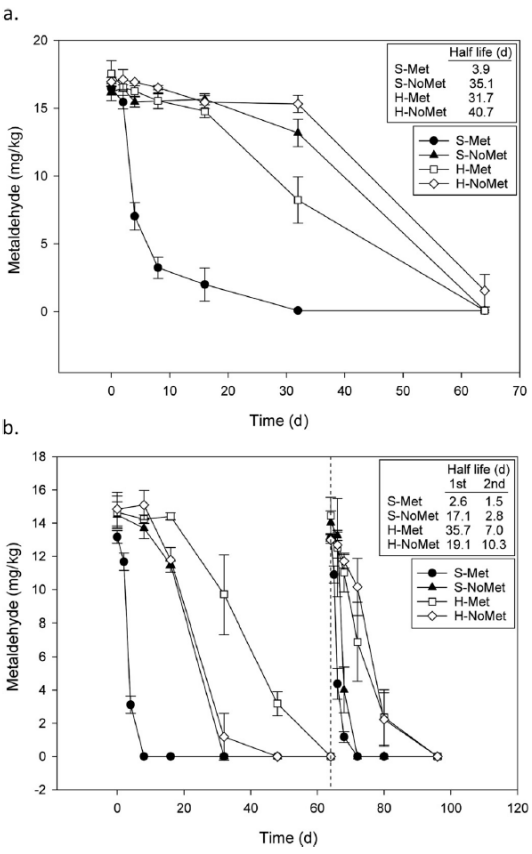


Fig. 2. **a**. Metaldehyde degradation profiles in freshly collected allotment soils after an initial application of 15 mg kg^{-1} soil. **b.** Degradation of metaldehyde allotment soils after storage (6 months) following an initial application of 15 mg kg^{-1} soil and a second application (dashed vertical line) at the same dose. Bars represent standard deviation for 3 replicates. Regression statistics are presented in Table S3. S-Met: Scarcroft Metaldehyde, S-NoMet: Scarcroft No Metaldehyde, H-Met: Hob Moor Metaldehyde, H-NoMet: Hob Moor No Metaldehyde.

metaldehyde. Hence, we hypothesized that metaldehyde-degrading organisms have been further enriched in soils after exposure to metaldehyde in the laboratory. Thus, to improve the recovery of metaldehyde-degrading isolates, selective enrichment for metaldehyde degraders was performed using the soil samples already incubated with metaldehyde for 64 d. This strategy permitted the isolation of metaldehyde-degrading strains from all four allotment soils in which the selective enrichment procedure had initially failed. Strains from diverse genera, including Gram-negative (*Acinetobacter*, *Pseudomonas*, *Sphingobium*, *Caballeronia*) and Gram-positive (*Rhodococcus*) isolates were successfully obtained. The identified metaldehyde-degrading strains are listed in Table 1. For all soils, in total, six Gram-negative and two Gram-positive strains were isolated. Whole-genome sequencing was performed for all the distinct taxa. Quality statistics for sequencing runs and assemblies are shown in Table S4.

The most abundant taxa in the final stage of these successful enrichment cultures, and their respective percentages in these original soils are also shown in Table S2. In the original soils, the genera with the highest abundances that could be confidently assigned taxonomy across all samples included the archaea *Candidatus* Nitrososphaera (9.6% of the initial community on average, SD: 3.1%) *Bacillus* (4.3%, SD: 2.5%) and *Kaistobacter* (2.4%, SD:1.7%), all which are commonly reported as some of the main genera found in soil (Zhalnina et al., 2013) and whose numbers decreased to undetectable levels at the end of the enrichment cultures. On the other hand, even though some genera such as *Acinetobacter*, *Sphingobium* and *Burkholderia*, all well-known xenobiotic degraders, were undetectable in the original soils with the sequencing depth used, their numbers increased to be amongst the dominant members of the enrichment cultures, which highlights both the strong selection pressure the populations were subjected to, and their ability to respond rapidly to it (Kurm et al., 2017).

3.4. Variation in the bacterial community due to metaldehyde addition and laboratory incubation

The addition of metaldehyde to the soil samples and the incubation in-soil increased the chances of isolating metaldehyde degraders. To further explore the variations in the microbial community occurring because of these procedures, changes were assessed by two different techniques: Most-probable number (MPN) of metaldehyde degraders and 16S rRNA amplicon sequencing of DNA from soil.

Fig. 3a shows the results of the MPN assay of metaldehyde-degrading bacteria at the end of the degradation assay. For all the allotment soils tested metaldehyde degraders were significantly more abundant on exposure to metaldehyde compared to untreated controls. For the soils except S-Met (which showed the fastest metaldehyde degradation even without metaldehyde exposure) this was a difference of several orders of magnitude.

The abundances of 16S rRNA gene copies were compared between soil samples during the metaldehyde degradation time course (Fig. 3b).

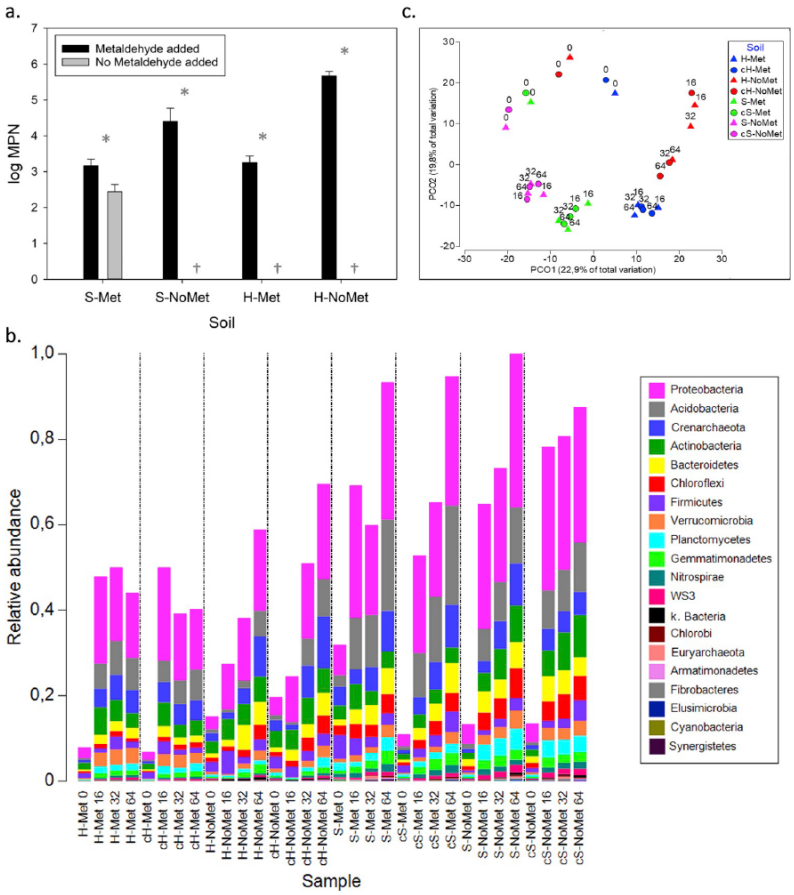


Fig. 3. a. Most probable number of culturable metaldehyde-degrading bacteria per g of soil 64 d after the addition of metaldehyde at 15 mg kg^{-1} or no addition at all. Dagger indicates <23 MPN g^{-1} soil. Bars represent standard deviation for 3 replicates. Asterisk indicates significant difference ($p < 0.05$). **b.** Abundance and classification by phyla of 16S rRNA gene copies in individual allotment soil samples (relative to the sample with the highest count) with and without an initial addition of metaldehyde (15 mg kg^{-1}) throughout a 64 d incubation period. Only the most abundant 20 phyla are listed in the key. **c.** Principal coordinates analysis plot of whole bacterial community 16S rRNA gene amplicons from individual allotment soil samples with and without an initial addition of metaldehyde (15 mg kg^{-1}) throughout a 64 d incubation period. Labels indicate day of sampling. Samples were rarefied to 56749 sequences each after denoising and quality control.

Quality statistics for sequencing runs are shown in Table S5. Irrespective of the addition of metaldehyde, there is a consistent marked increase in bacterial abundance throughout the incubation period for all soils, which may also increase the chances of successful isolation of metaldehyde degraders. The fraction of the population at the Phylum level corresponding to Proteobacteria showed an increase for all soils between the 0 d and 16 d time points and showed a slow decline thereafter.

To explore the differences in the composition of bacterial communities during the in-soil metaldehyde degradation assay, a PCO ordination of whole bacterial community 16S rRNA gene amplicons was constructed (Fig. 3c). Principal coordinates 1 and 2 accounted for 42.7% of the variation. PERMANOVA analyses revealed that soil of origin (Pseudo-F = 6.26, p = 0.0001, df = 31) and time (Pseudo-F = 2.6058, p = 0.0001, df = 31) but not metaldehyde addition in the laboratory (Pseudo-F = 0.26951, p = 0.9999, df = 31) significantly influenced the soil community composition. Sample group dispersions were homogeneous for all factors as revealed by PERMDISP analyses.

3.5. Growth of isolated strains using metaldehyde as a sole source of carbon

Growth of all the isolated degrading strains (and the previously isolated *A. calcoaceticus* E1) on metaldehyde (150 mg L^{-1}) as the sole source of carbon was achieved, and a strong correlation with the disappearance of metaldehyde was observed, supporting the conclusion that they use metaldehyde as a carbon and energy source (Fig. 4, Table S6). *A. bohemicus* JMET-C stood out as the strain with the shortest lag phase, doubling time, time required for metaldehyde removal below the limit of detection and the highest maximum compound degradation rate (Fig. 4). The rest of the *Acinetobacter* strains followed with respect to the time needed for metaldehyde elimination; *P. vancouverensis* SMET-B came next. The *Sphingobium* CMET-H/HMET-G strains were intermediate in this regard, while the *Rhodococcus globerulus* HMET-A/HNO-A and *C. jiangsuensis* SNO-D strains were slowest.

3.6. Comparative genomic analysis of metaldehyde-degrading strains

Diverse metaldehyde-degrading strains were isolated, and a draft genome for each one was obtained. Comparative genomic analyses focused on identifying shared genes encoding proteins needed for metaldehyde degradation.

A previous study from our group had identified an *Acinetobacter* strain (E1) capable of degrading metaldehyde (Thomas et al., 2017); additionally, two other strains of *Acinetobacter* (*A. bohemicus* JMET-C and *A. lwoffii* SMET-C) were isolated in this study. The predicted proteomes from these strains, along with those of the

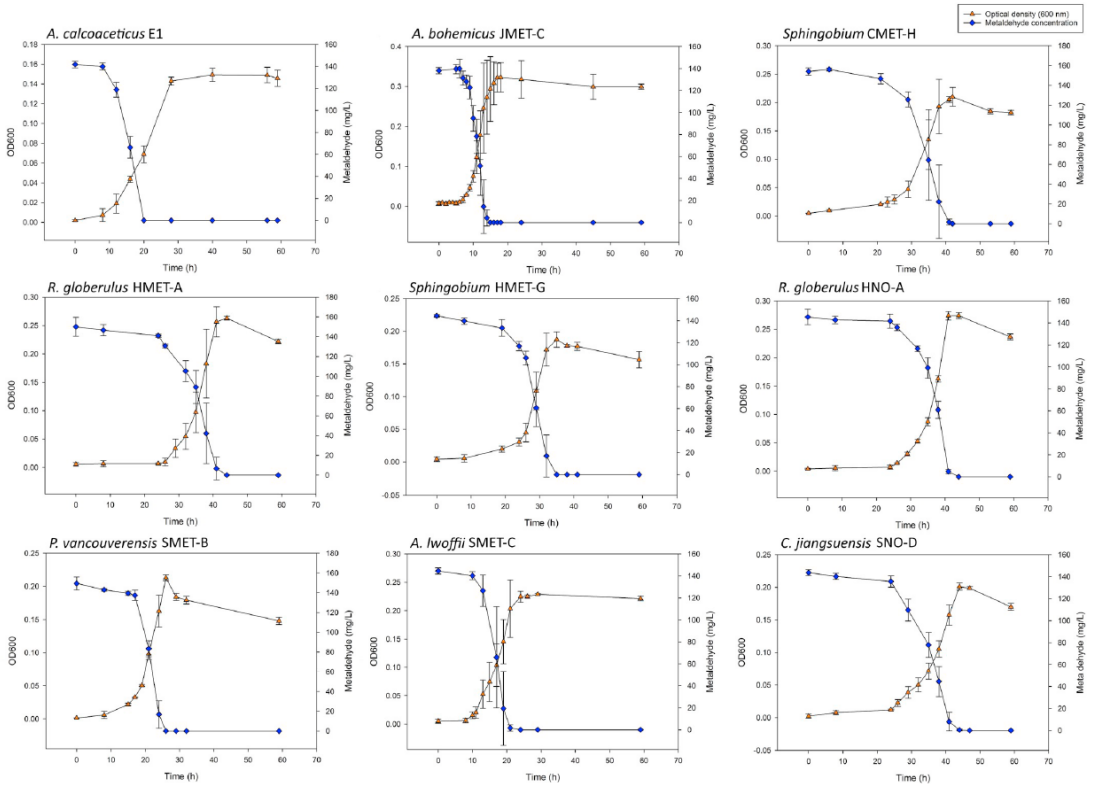


Fig. 4. Growth of degrading strains on metaldehyde (150 mg L^{-1}) as the sole source of carbon. The inoculum was substituted by an equal volume MSM in the abiotic controls. Bars represent standard deviation for 3 replicates. Final metaldehyde concentration in the abiotic controls was 99.9% (CV = 3.0%) of the starting concentration.

metaldehyde-degrading strains *P. vancouverensis* SMET-B, *C. jiangsuensis* SNO-D, and the non-degrading strain *A. calcoaceticus* RUH2202 were compared through Blast Score Ratio (BSR) analysis using *A. calcoaceticus* E1 as reference. A total of 65 candidate proteins were identified as present in all metaldehyde-degrading strains of *Acinetobacter* but absent from *A. calcoaceticus* RUH2202. Inclusion of the genomes from more evolutionarily divergent strains SNO-D and SMET-B decreased the number of predicted proteins shared amongst the metaldehyde-degrading strains and absent from RUH2202 to four (Table 2). Identical results were obtained when the analysis was repeated using *A. bohemicus* ANC3994 as non-degrading strain. A graphical representation of the results is displayed in Fig. 5.

The identified proteins corresponded to a single cluster of four apparently horizontally-transferred genes. Predicted protein sequences were used to search NCBI's conserved domain database (Marchler-Bauer et al., 2015), using the cluster from *A. bohemicus* JMET-C as query. Thresholds of 50% and 70% identity have been proposed for assignment of third and full Enzyme Commission numbers at the domain level (Addou et al., 2009), and these values were used to support predicted functional assignment and naming of genes. Where only lower sequence identity to existing genes was observed, genes were denoted *mah* (for *met*al*de*hyde). The first gene in the cluster (*mahX*) contains a main

Table 2
Predicted proteins shared between metaldehyde-degrading strains, absent from non-degrading strains and their respective BSR values when compared against the reference genome (*A. calcoaceticus* E1).

	MahX		MahY		AldH		TnpA	
	aa length	BSR[a]	aa length	BSR[a]	aa length	BSR[a]	aa length	BSR[a]
A. calcoaceticus E1	314	1.000	149	1.000	231	1.000	503	1.000
A. calcoaceticus RUH2202	NP[b]	-	NP[b]	-	495	0.441	NP[b]	-
A. bohemicus JMET- C	314	1.000	149	1.000	231	1.000	503	1.000
A. lwoffii SMET-C	314	0.994	149	0.993	231	1.000	387	0.997
P. vancouverensis SMET-B	314	0.975	149	0.993	231	0.995	327	0.997
C. jiangsuensis SNO-D	314	0.984	149	0.993	88	1.000	262	1.000

[a] BSR: Blast-Score Ratio.
[b] NP: not present.

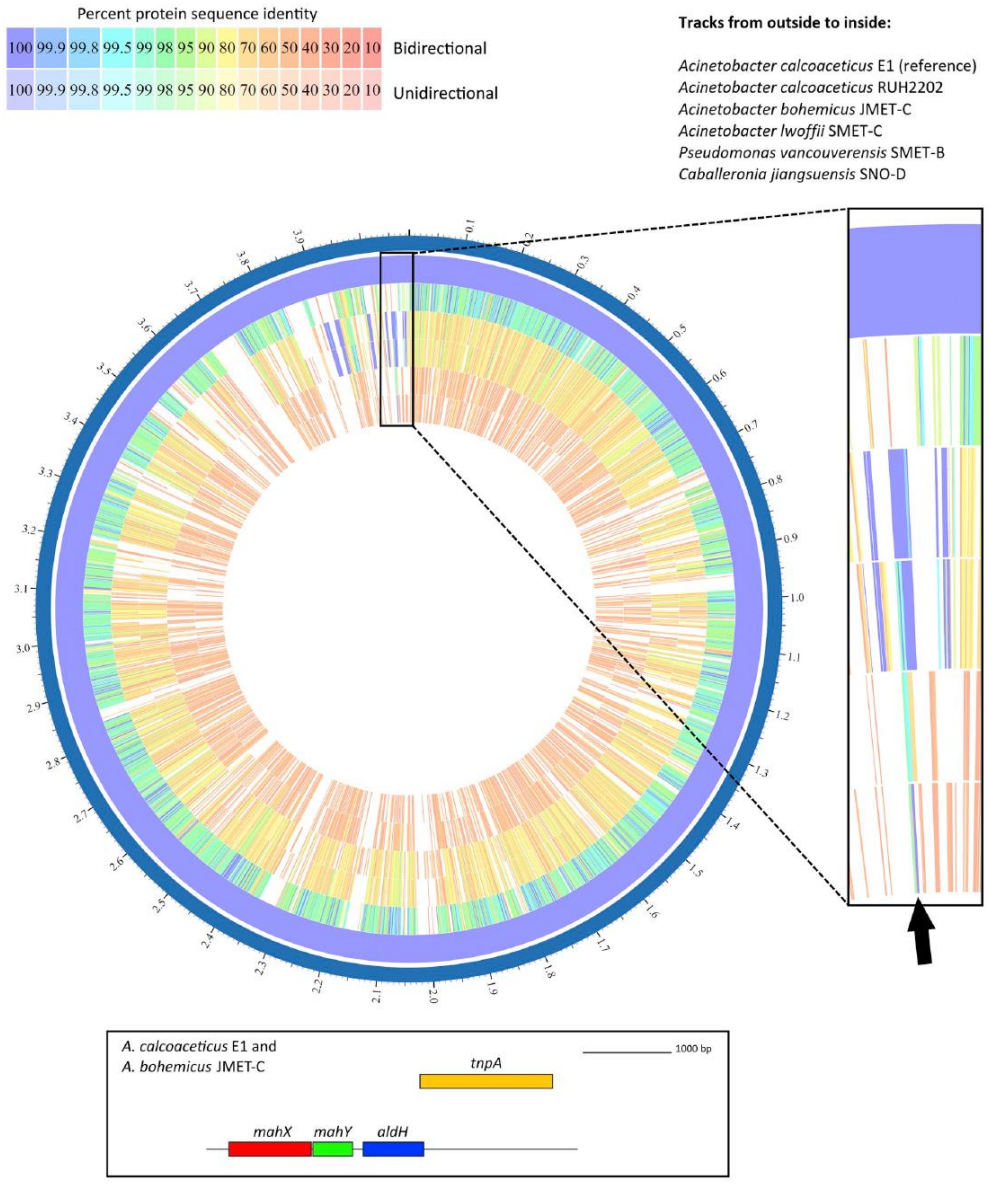


Fig. 5. Proteome comparison between metaldehyde-degrading and non-degrading strains using *Acinetobacter calcoaceticus* E1 as reference. The analysis was performed using PATRIC Proteome Comparison Service (Wattam et al., 2017). A diagram of the shared metaldehyde-degrading gene cluster is shown in the inset. *mahX*: 2-oxoglutarate and Fe(II)-dependent oxygenase; *mahY*: vicinal oxygen chelate protein; *aldH*: NAD(P)$^+$-dependent aldehyde dehydrogenase; *tnpA*: Y2-transposase.

domain classified into the 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily, with the highest sequence identity (49%) to a protein involved in biosynthesis of mitomycin antibiotics from *Sphingobium japonicum*. The second gene (*mahY*) is related to the vicinal oxygen chelate family, which is found in a variety of structurally related metalloproteins, including the type I extradiol dioxygenases, glyoxalase I and a group of antibiotic resistance proteins, with the highest sequence identity (41%) to a hypothetical unannotated protein from *Mycolicibacterium moriokaense*. The third gene (*aldH*) could be confidently assigned as an NAD(P)$^+$-dependent aldehyde dehydrogenase due to high identity (78%) to an aldehyde dehydrogenase from *Solimonas* sp. The fourth gene (*tnpA*), annotated as Y2-transposase, is almost identical (99%) to a transposase from *Pseudomonas pseudoalcaligenes*. All these functions are consistent with a potential horizontally-transferable metaldehyde degradation pathway. In *P. vancouverensis* SMET-B the *tnpA* was found to be located elsewhere in the genome. In *C. jiangsuensis* SNO-D, *aldH* was truncated.

A possible horizontal gene transfer event to the *Acinetobacter* taxon

was supported by the presence of the transposase gene in the cluster, IS6 family insertion sequences at both ends of the respective contigs and different GC nucleotide percentages between the gene cluster (59.81–62.74%) and the draft whole genome sequences of metaldehyde-degrading *Acinetobacter* strains (38.7–40.2%), all parametric indicators of horizontal gene transfer (Ravenhall et al., 2015).

Chemical mutagenesis of *A. calcoaceticus* E1 led to isolation of four strains unable to degrade metaldehyde. In each case, the deletion of part of the gene encoding the putative oxygenase (*mahX*) was the only shared difference from the wild-type (Fig. S1), which further supports the involvement of the gene cluster in metaldehyde degradation.

3.7. Heterologous expression of mahX confers ability to degrade metaldehyde on E. coli

mahX, *mahX* + *mahY*, and *mahX* + *mahY* + *aldH* were cloned into pBR322 and transformed into *E. coli*. Fig. 6 illustrates that *mahX* was necessary and sufficient to confer the ability to degrade metaldehyde on *E. coli*, confirming the inferences from comparative genomics and chemical mutagenesis. The rate of metaldehyde degradation is faster in *E. coli* containing *mahX* only, than in strains also expressing *mahY*. The reasons for this are unknown, but may be related to difference in gene expression between these different constructs. *E. coli* bearing *mahY* (under the *mahX* promoter) cloned into pBR322 is unable to degrade metaldehyde (data not shown).

4. Discussion

Previous work has indicated microbial activity is involved in degradation of the xenobiotic pesticide metaldehyde (Simms et al., 2006; Thomas et al., 2017). Here we have used a systematic molecular and microbiological approach to analyse and improve isolation techniques, and gain insight into the abundance, distribution and mechanisms of microbial metaldehyde degradation.

We successfully combined the isolation of diverse metaldehyde-degrading strains with increasingly affordable whole-genome sequencing to, by means of comparative genomics, identify a xenobiotic-degrading gene cluster. This strategy has proven effective for other xenobiotics in the past as well (Yan et al., 2016), and as whole-genome sequencing becomes increasingly affordable, it has the potential of becoming a very important approach to identifying organic compound-degrading gene clusters. This functional assignment was subsequently confirmed by chemical mutagenesis and heterogeneous expression of the enzyme activity. This illustrates the value of obtaining a broad collection of isolates for the identification of degrading mechanisms.

Availability of labile carbon is considered the main limiting factor for microbial growth in soil (Aldén et al., 2001), and metaldehyde, when added to the soil, constitutes such a source of carbon. In this context, the presence of genes for its degradation provide the host with the ability to utilize this readily available carbon source, constituting a selective advantage. Genes *mahX* and *mahY* have a moderate similarity to other well-characterized genes, so they may share an evolutionary ancestor with them. However, they appear to have diverged sufficiently so that new catalytic properties, such as substrate specificity towards metaldehyde, has emerged, was selected for, and transferred to other hosts via transposable elements in plasmids or other vectors. *mahX* encodes a predicted 2-oxoglutarate-dependent oxygenase. Given that this gene is sufficient to bring about at least the initial step of metaldehyde degradation, we propose that MahX is an oxygenase that activates metaldehyde metabolism by oxygenation and ring cleavage. The predicted product is a hemiacetal (1,3,5,7-tetramethyl-2,4,6-trioxa-1-hydroxy-7-octanone) (Fig. 7). Whilst this hemiacetal is unstable, the timescale of its chemical degradation is likely to be minutes to hours (Chiang and Kresge, 1985). We speculate that MahY acts as a lyase that accelerates the iterative breakdown of the hemiacetal intermediate to acetaldehyde (Fig. 7). MahY is most closely related to the vicinal oxygen chelate (VOC) superfamily (He and Moran, 2011). Whilst the predicted substrate here is not a VOC *per se*, VOC family members bind substrates with two oxygen atoms and some members of the family are lyases in keeping with the predicted function here. Acetaldehyde generated from MahY is predicted to be converted to acetate by AldH (Fig. 7), and subsequently incorporated into central metabolism.

Since the degrading gene cluster identified in *Acinetobacter*, *Caballeronia* and *Pseudomonas* is not present in some of the other metaldehyde-degrading isolates from this study (*Sphingobium* sp. CMET-H and HMET-G, *R. globerulus* HMET-A and HNO-A) and from an earlier study (*Variovorax* sp. E3) (Thomas et al., 2017), as shown by whole-genome sequencing and PCR-based gene detection (data not shown), it is clear that additional metaldehyde-degrading mechanisms are found in nature.

Previous attempts to isolate metaldehyde-degrading strains from soils resulted in varying degrees of success. Thomas and collaborators (Thomas et al., 2017) isolated two metaldehyde-degrading strains from

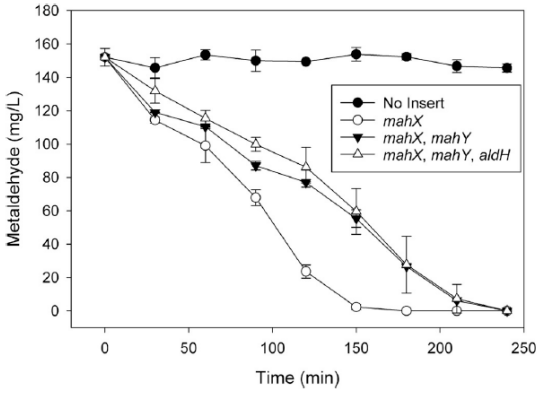


Fig. 6. *mahX* is sufficient to confer degradation of metaldehyde. Degradation of metaldehyde by *E. coli* was measured following growth in LB liquid media. *E. coli* DH5α carried plasmid pBR322 (solid circles), and derivatives of this plasmid containing *mahX* (open circles), *mahX* & *mahY* (filled triangles), or *mahX*, *mahY* and *aldH* (open triangles).

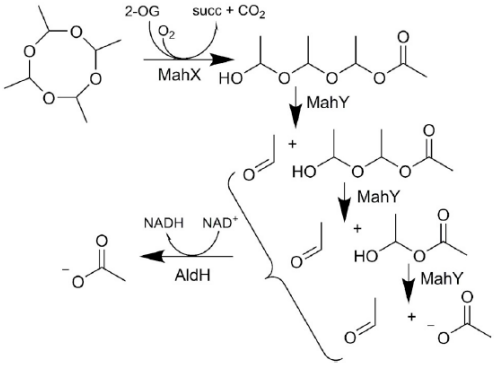


Fig. 7. Predicted pathway for metaldehyde degradation. MahX is related to 2-oxoglutarate (2-OG)-dependent oxygenases that generate succinate (succ) and CO_2. MahX oxygenates metaldehyde to release a linear hemiacetal that is cleaved iteratively into acetaldehyde + a shorter chain hemiacetal, and eventually acetate. AldH oxidises acetaldehyde to acetate in an NAD^+-dependent reaction.

domestic soils, nevertheless, additional attempts using three different previously-exposed agricultural soils failed. Initially, we obtained similar results; we could only sporadically isolate degraders from previously-exposed soils. Even though previous applications of metaldehyde to the soil in the field seemed to aid in the subsequent isolation of degraders, it was far from a guarantee. Many variables such as the length of time between application and soil sampling, the dose of pesticide, its formulation (and thus distribution and fate in the field), and the sampling regime, may all influence the abundance of metaldehyde degraders in the soil samples, and thus the success of isolation strategies. Furthermore, laboratory culture conditions cannot fully replicate the ideal conditions to suit the physiology of many of members of the soil microbial community.

The fact that metaldehyde is normally applied as a pellet in the field (vs. liquid spray forms) influences its behaviour and distribution in soil (Bond, 2018). This is important because a gradient of decreasing concentrations in soil would be expected as distance from the pellet increases. The input of labile C sources to soil increases the abundance and activity of microorganisms generating microbial hotspots (Kuzyakov and Blagodatskaya, 2015). The distribution of metaldehyde after pellet applications may generate defined zones of enhanced metaldehyde degradation in soil. This would pose a challenge when sampling, because high biodegradation hotspots, with elevated numbers of degraders, could easily be missed.

We were able to consistently isolate metaldehyde degraders from horticultural soils, regardless of their metaldehyde exposure history, after adding metaldehyde to soil samples in the laboratory and incubating them for a defined period. This approach minimized the challenges to isolation that would arise from distribution of activity in hotspots. Further analysis revealed that metaldehyde addition increased the number of culturable metaldehyde degraders in the samples, while incubation in laboratory conditions (initial homogenization, 25 °C, 100% relative humidity) also increased the total bacterial biomass, greatly facilitating, in combination, the isolation of degraders. Even though this approach has been used in the past for other pesticides (Goda et al., 2010; Perruchon et al., 2016; Singh et al., 2004) it seems to be of particular importance for the isolation of metaldehyde degraders because the uneven distribution of the pesticide generated from normal field applications makes the direct isolation from freshly collected soils difficult.

In this study we observed that metaldehyde degradation occurred faster in soils previously exposed to metaldehyde in the field compared to soils with very similar physicochemical characteristics but not directly exposed to the pesticide. Furthermore, even a single metaldehyde addition to soil samples in the laboratory led to accelerated degradation of metaldehyde and accumulation of culturable metaldehyde-degrading microbes. This illustrates that metaldehyde degrading strains are strongly selected by metaldehyde in soil.

Culture-independent analysis during metaldehyde degradation showed that the main factors governing bacterial community composition were soil origin and incubation time, not metaldehyde. Several studies have found important pesticide-driven changes in community structure with other pesticides only when using doses of several times the recommended application rate or after repeated additions (Crouzet et al., 2010; Cycoń et al., 2013; Itoh et al., 2014). Nevertheless, at a finer level, several specific taxa were enriched after metaldehyde application (Table S7). Notably, the lists of most enriched taxa in all soils after metaldehyde application were dominated by anaerobes, which may not reflect the specific groups executing metaldehyde degradation in soil but may instead be related to increased bacterial metabolism and abundance, leading to anaerobic conditions developing in the soils during the incubation due to increased oxygen respiration (Shennan et al., 2018; Streminska et al., 2014). This is consistent with the fact that enrichment of specific anaerobic taxa was observed even in the already metaldehyde-adapted soil S-Met.

A metaldehyde-degrading *Sphingobium* was isolated following enrichment (in soil H-Met), and this genus was identified, through amplicon sequencing, as being enriched > 300-fold by metaldehyde in the soil (Table S7). Beyond that, no other cultured metaldehyde-degraders were identified as enriched in soil amplicon sequencing. This is presumably a result of metaldehyde degradation being a rare trait, possessed by a relatively low proportion of the total microbial community (reflected by 10^3-10^5 metaldehyde-degraders (Fig. 3a) out of an estimated 10^8-10^9 microbial cells/g of soil in our samples).

This study highlights the continuing value of "traditional" experimental enrichment methods, and their application in coordination with contemporary molecular methods. Here we have provided insight into the diversity and mechanism of biological metaldehyde degradation; this understanding will be essential for predicting the environmental fate of the compound and for optimizing and monitoring the performance of engineered biological systems for metaldehyde removal from drinking water.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

VCG was supported by the University of Costa Rica and the University of York. JM, SJ and CS are grateful to the Natural Environment Research Council (NERC) and Thames Water for grant NE/N009061/1. JM, TH and CS are grateful for the BBSRC studentship that supported JCT.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.soilbio.2019.107702.

References

Abraham, J., Silambarasan, S., 2013. Biodegradation of chlorpyrifos and its hydrolyzing metabolite 3,5,6-trichloro-2-pyridinol by Sphingobacterium sp. JAS3. Process Biochemistry 48, 1559–1564. https://doi.org/10.1016/J.PROCBIO.2013.06.034.

Addou, S., Rentzsch, R., Lee, D., Orengo, C.A., 2009. Domain-based and family-specific sequence identity thresholds increase the levels of reliable protein function transfer. Journal of Molecular Biology 387, 416–430. https://doi.org/10.1016/J.JMB.2008.12.045.

Aldén, L., Demoling, F., Bååth, E., 2001. Rapid method of determining factors limiting bacterial growth in soil. Applied and Environmental Microbiology 67, 1830–1838. https://doi.org/10.1128/AEM.67.4.1830-1838.2001.

Arbeli, Z., Fuentes, C.L., 2007. Accelerated biodegradation of pesticides: an overview of the phenomenon, its basis and possible solutions; and a discussion on the tropical dimension. Crop Protection 26, 1733–1746. https://doi.org/10.1016/J.CROPRO.2007.03.009.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V. M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology: A Journal of Computational Molecular Cell Biology 19, 455–477. https://doi.org/10.1089/cmb.2012.0021.

Bond, S.G., 2018. Assessing the Fate of Metaldehyde Applied to Arable Soils. University of Leicester.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7, 335–336. https://doi.org/10.1038/nmeth.f.303.

Carpenter, M., 1989. Metaldehyde Draft Assessment Report, vol. 3. B8.

Castle, G.D., Mills, G.A., Gravell, A., Jones, L., Townsend, I., Cameron, D.G., Fones, G.R., 2017. Review of the molluscicide metaldehyde in the environment. Environmental Sciences: Water Research & Technology 3, 415–428. https://doi.org/10.1039/C7EW00039A.

Chiang, Y., Kresge, A.J., 1985. Kinetics of hydrolysis of acetaldehyde ethyl hemiacetal in aqueous solution. Journal of Organic Chemistry 50, 5038–5040. https://doi.org/10.1021/jo00225a007.
Chief Inspector of Drinking Water, 2017. Drinking Water 2016 – Public Water Supplies for England and Wales, Quarter 3, July-September 2016. London.
Cranor, W., 1990. Metaldehyde Draft Assessment Report, vol. 3. B8.
Crouzet, O., Batisson, I., Besse-Hoggan, P., Bonnemoy, F., Bardot, C., Poly, F., Bohatier, J., Mallet, C., 2010. Response of soil microbial communities to the herbicide mesotrione: a dose-effect microcosm approach. Soil Biology and Biochemistry 42, 193–202. https://doi.org/10.1016/J.SOILBIO.2009.10.016.
Cycoń, M., Markowicz, A., Borymski, S., Wójcik, M., Piotrowska-Seget, Z., 2013. Imidacloprid induces changes in the structure, genetic diversity and catabolic activity of soil microbial communities. Journal of Environmental Management 131, 55–65. https://doi.org/10.1016/J.JENVMAN.2013.09.041.
Dinamarca, M.A., Cereceda-Balic, F., Fadic, X., Seeger, M., 2007. Analysis of s-triazine-degrading microbial communities in soils using most-probable-number enumeration and tetrazolium-salt detection. International Microbiology: The Official Journal of the Spanish Society for Microbiology 10, 209–215.
Dobritsa, A.P., Samadpour, M., 2016. Transfer of eleven species of the genus Burkholderia to the genus Paraburkholderia and proposal of Caballeronia gen. nov. to accommodate twelve species of the genera Burkholderia and Paraburkholderia. International Journal of Systematic and Evolutionary Microbiology 66, 2836–2846. https://doi.org/10.1099/ijsem.0.001065.
Eckert, M., Fleischmann, G., Reinhard, J., Bolt, H., Golka, K., 2012. Acetaldehyde. In: Ullmann's Encyclopedia of Industrial Chemistry.
European Food Safety Authority, 2010. Conclusion on the peer review of the pesticide risk assessment of the active substance metaldehyde. EFSA Journal 8, 1856. https://doi.org/10.2903/j.efsa.2010.1856.
Fierer, N., 2017. Embracing the unknown: disentangling the complexities of the soil microbiome. Nature Reviews Microbiology 15, 579–590. https://doi.org/10.1038/nrmicro.2017.87.
FOCUS, 2006. Guidance Document on Estimating Persistence and Degradation Kinetics from Environmental Fate Studies on Pesticides in EU Registration.
Garthwaite, D., Barker, I., Ridley, L., Mace, A., Parrish, G., MacArthur, R., Lu, Y., 2018. Pesticide Usage Survey Report 271 - Arable Crops in the United Kingdom (Version 2) 2016. Sand Hutton.
Geißdörfer, W., Kok, R.G., Ratajczak, A., Hellingwerf, K.J., Hillen, W., 1999. The genes rubA and rubB for alkane degradation in Acinetobacter sp. strain ADP1 are in an operon with estB, encoding an esterase, and oxyR. Journal of Bacteriology 181, 4292–4298.
Gimingham, C.T., 1940. Some recent contributions by English workers to the development of methods of insect control. Annals of Applied Biology 27, 161–175. https://doi.org/10.1111/j.1744-7348.1940.tb07486.x.
Goda, S.K., Elsayed, I.E., Khodair, T.A., El-Sayed, W., Mohamed, M.E., 2010. Screening for and isolation and identification of malathion-degrading bacteria: cloning and sequencing a gene that potentially encodes the malathion-degrading enzyme, carboxylestrase in soil bacteria. Biodegradation 21, 903–913. https://doi.org/10.1007/s10532-010-9350-3.
Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29, 1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
He, P., Moran, G.R., 2011. Structural and mechanistic comparisons of the metal-binding members of the vicinal oxygen chelate (VOC) superfamily. Journal of Inorganic Biochemistry 105, 1259–1272. https://doi.org/10.1016/J.JINORGBIO.2011.06.006.
Itoh, H., Navarro, R., Takeshita, K., Tago, K., Hayatsu, M., Hori, T., Kikuchi, Y., 2014. Bacterial population succession and adaptation affected by insecticide application and soil spraying history. Frontiers in Microbiology 5, 457. https://doi.org/10.3389/fmicb.2014.00457.
Kegley, S.E., Hill, B.R., Orme, S., Choi, A.H., 2016. PAN pesticide database, pesticide action network, North America [WWW Document]. http://www.pesticideinfo.org/ accessed 4.30.18.
Kurm, V., Van Der Putten, W.H., De Boer, W., Naus-Wiezer, S., Gera Hol, W.H., 2017. Low abundant soil bacteria can be metabolically versatile and fast growing. Ecology 98, 555–564. https://doi.org/10.1002/ecy.1670.
Kuzyakov, Y., Blagodatskaya, E., 2015. Microbial hotspots and hot moments in soil: concept & review. Soil Biology and Biochemistry 83, 184–199. https://doi.org/10.1016/J.SOILBIO.2015.01.025.
Lazartigues, A., Thomas, M., Cren-Olivé, C., Brun-Bellut, J., Le Roux, Y., Banas, D., Feidt, C., 2013. Pesticide pressure and fish farming in barrage pond in Northeastern France. Part II: residues of 13 pesticides in water, sediments, edible fish and their relationships. Environmental Science and Pollution Research 20, 117–125. https://doi.org/10.1007/s11356-012-1167-7.
Lewis, K.A., Tzilivakis, J., Warner, D.J., Green, A., 2016. An international database for pesticide risk assessments and management. Human and Ecological Risk Assessment: An International Journal 22, 1050–1064. https://doi.org/10.1080/10807039.2015.1133242.
Maheshwari, M., Abulreesh, H.H., Khan, M.S., Ahmad, I., Pichtel, J., 2017. Horizontal gene transfer in soil and the rhizosphere: impact on ecological fitness of bacteria. In: Agriculturally Important Microbes for Sustainable Agriculture. Springer Singapore, Singapore, pp. 111–130. https://doi.org/10.1007/978-981-10-5589-8_6.
Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Bryant, S.H., 2015. CDD: NCBI's conserved domain database. Nucleic Acids Research 43, D222–D226. https://doi.org/10.1093/nar/gku1221.
Möllerfeld, J., Römbke, J., Heller, M., 1993. Metaldehyde Draft Assessment Report, vol. 3. B8.
Neilson, A.H., Allard, A.-S., 2012. Organic Chemicals in the Environment, Organic Chemicals in the Environment. CRC Press. https://doi.org/10.1201/b12492.
Perruchon, C., Patsioura, V., Vasileiadis, S., Karpouzas, D.G., 2016. Isolation and characterisation of a *Sphingomonas* strain able to degrade the fungicide *ortho*-phenylphenol. Pest Management Science 72, 113–124. https://doi.org/10.1002/ps.3970.
Rasko, D.A., Myers, G.S., Ravel, J., 2005. Visualization of comparative genomic analyses by BLAST score ratio. BMC Bioinformatics 6, 2. https://doi.org/10.1186/1471-2105-6-2.
Ravenhall, M., Škunca, N., Lassalle, F., Dessimoz, C., 2015. Inferring horizontal gene transfer. PLoS Computational Biology 11, 1–16. https://doi.org/10.1371/journal.pcbi.1004095.
Rolph, C., 2016. 21st Century Biological Processes for Metaldehyde Removal in Drinking Water Treatment. Cranfield University.
Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
Shennan, C., Muramoto, J., Koike, S., Baird, G., Fennimore, S., Samtani, J., Bolda, M., Dara, S., Daugovish, O., Lazarovits, G., Butler, D., Rosskopf, E., Kokalis-Burelle, N., Klonsky, K., Mazzola, M., 2018. Anaerobic soil disinfestation is an alternative to soil fumigation for control of some soilborne pathogens in strawberry production. Plant Pathology 67, 51–66. https://doi.org/10.1111/ppa.12721.
Simms, L.C., Dawson, J.J.C., Paton, G.I., Wilson, M.J., 2006. Identification of environmental factors limiting plant uptake of metaldehyde seed treatments under field conditions. Journal of Agricultural and Food Chemistry. https://doi.org/10.1021/jf060231a.
Singh, B.K., Walker, A., Morgan, J.A.W., Wright, D.J., 2004. Biodegradation of chlorpyrifos by Enterobacter strain B-14 and its use in bioremediation of contaminated soils. Applied and Environmental Microbiology. https://doi.org/10.1128/AEM.70.8.4855-4863.2004.
Smets, W., Leff, J.W., Bradford, M.A., McCulley, R.L., Lebeer, S., Fierer, N., 2016. A method for simultaneous measurement of soil bacterial abundances and community composition via 16S rRNA gene sequencing. Soil Biology and Biochemistry 96, 145–151. https://doi.org/10.1016/J.SOILBIO.2016.02.003.
Streminska, M.A., van der Wurff, A.W.G., Runia, W.T., Thoden, T.C., Termorshuizen, A. J., Feil, H., 2014. Changes in bacterial and fungal abundance in soil during the process of anaerobic soil disinfestation. Acta Horticulturae 95–102. https://doi.org/10.17660/ActaHortic.2014.1041.9.
Tao, B., Fletcher, A.J., 2013. Metaldehyde removal from aqueous solution by adsorption and ion exchange mechanisms onto activated carbon and polymeric sorbents. Journal of Hazardous Materials 244 (245), 240–250. https://doi.org/10.1016/J.JHAZMAT.2012.11.014.
Thomas, J.C., Helgason, T., Sinclair, C.J., Moir, J.W.B., 2017. Isolation and characterization of metaldehyde-degrading bacteria from domestic soils. Microbial Biotechnology 10, 1824–1829. https://doi.org/10.1111/1751-7915.12719.
Vishniac, W., Santer, M., 1957. The thiobacilli. Bacteriological Reviews 21, 195–213.
Wattam, A.R., Davis, J.J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., Conrad, N., Dietrich, E.M., Disz, T., Gabbard, J.L., Gerdes, S., Henry, C.S., Kenyon, R.W., Machi, D., Mao, C., Nordberg, E.K., Olsen, G.J., Murphy-Olson, D.E., Olson, R., Overbeek, R., Parrello, B., Pusch, G.D., Shukla, M., Vonstein, V., Warren, A., Xia, F., Yoo, H., Stevens, R.L., 2017. Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. Nucleic Acids Research 45, D535–D542. https://doi.org/10.1093/nar/gkw1017.
Yan, X., Gu, T., Yi, Z., Huang, J., Liu, X., Zhang, J., Xu, X., Xin, Z., Hong, Q., He, J., Spain, J.C., Li, S., Jiang, J., 2016. Comparative genomic analysis of isoproturon-mineralizing sphingomonads reveals the isoproturon catabolic mechanism. Environmental Microbiology 18, 4888–4906. https://doi.org/10.1111/1462-2920.13413.
Zhalnina, K., de Quadros, P.D., Gano, K.A., Davis-Richardson, A., Fagen, J.R., Brown, C. T., Giongo, A., Drew, J.C., Sayavedra-Soto, L.A., Arp, D.J., Camargo, F.A.O., Daroub, S.H., Clark, I.M., McGrath, S.P., Hirsch, P.R., Triplett, E.W., 2013. Ca. Nitrososphaera and Bradyrhizobium are inversely correlated and related to agricultural practices in long-term field experiments. Frontiers in Microbiology 4. https://doi.org/10.3389/fmicb.2013.00104.
Zhang, H., Wang, C., Lu, H., Guan, W., Ma, Y., 2011. Residues and dissipation dynamics of molluscicide metaldehyde in cabbage and soil. Ecotoxicology and Environmental Safety 74, 1653–1658. https://doi.org/10.1016/J.ECOENV.2011.05.004.
Zhang, X., Dai, X., 2006. Degradation and determination of the residue of metaldehyde in tobacco and soil. Chinese Journal of Pesticide Science 4, 344–348.

Annex 2: Specific gene sequences – Metaldehyde-degrading gene cluster

>mahX_AcalcoaceticusE1 (945 bp)
atgaagcaagagcttgagtccgccgtgaatgcgcgagtgatcgagtcgaacaagctcgatggat
cgagcctgatcggtcaggagcgtcagccggacactgatttttacgctgccctctccgacatggg
gctcggcgggcacgtggcccacttggatcagtacggctattgcgtggttcctccgtccgacttc
gacgatctcgggttgacagcggaggccaaaagaaaagttttggagatcgcggaaagacgcagcg
gtatccggcctgacagcgaaacgggtgctacacactcgatcggcgaatcggcggtcgggcagtg
tatgcactacctgctcttcgaagaccccgtcttcgagaagatgctcattcatccggtggtattg
gcgttccaccgttatctcctcggcaggagcggaaggctgagcgccatgagcgcgatgcttcggg
gcccgggtacgcccgccttagctacgcatgcggacttggtgatggtgccgccgccttggccgat
gttcgcgcaggtatgcaatatctcttgggcattgacggactacacgaaggaaaatggcgctacc
gcgatcgttcctggaagtcacaagctctgccggccaccgacggacgccgagatcgcggatacca
gtaagctgatcgcggtaaccgcgccggctggatccctcgtcatttggcacggcaacacttggca
tggctccttcgcgaaggtgagcccgggactgcgtatgcagatcatcatgtacatgtgccgcact
cacgtgatgcctcaggaatggtatctcgacaaggtaacgccggagatactgcagagaaatggaa
ccgagttcgccgaaatgctcgggatcggtcatccgtatccgtttgctccttcgggtccggattg
gcaaaacgtcgttcgggcgttcggccaggcaacgacgttgaacggctaa

>mahY_AcalcoaceticusE1 (450 bp)
atgaacaaattacatcgagtcgttgtcgcggttaaggatctcgacgaagctgcaggtcggtacg
agcggatcttcgccgtccctttcgtccgcaccggcccctatgtcgcgtccatgggcgtgaaggt
cgcgggcgcgtgggggctcggtgtcgaacttattcagccgatgcccggcagcgacagccagttt
gctcaggacatccagaggcacctgaacgagcgcggggaagggctgtacggggtggtgtttcaga
ctcgtacgatgaaatctgacatcgagcacctggagaaaaatgcgttcgtagcatacggaccgac
cttctctttctccagcagcgttctggaaaccgagttcggtggcgccttcagccggttcgaggag
acggtatttaccccggagcgcctcggctatctggtggcggcaatggacgcttccccaacacgat
aa

>aldH_AcalcoaceticusE1 (696 bp)
atggacggagatcgtggtttgacggcattgctgccgtcggtgaaggaatttctgggttcgccgc
gcaaggcattcatcgacggcaaatgggttgctgcgaagagcggcaagacgttcgaggtgttcaa
cccggcgacgggcagcgtgatcggacatgcggcggcgtgtgagaaggcggacgtggacgaggcg
gtgaaggctgcgcgcaaggcgttcgacacgggtccctggaccaagatgacgccgtccgaacgcg
gtcggatcatctggaaggtgggcgatctgatcctgaagtacaccgaggaactcgcgcagctcga
gtcgatcgacaacggcaagccgatctcgatcgcgcgcgctgccgacgtcgtgctcgctgccgac
atcttccattacatggccggctgggccacaaagatcgagggcagttcgctctcgctgtcggttc
cctacaccccgggcgtcgagtaccaggcgttcacgcgcaaggagccggtcggcgtcgtcgcgca
gatcattccctggaacttcccgctgctgatggccgcgtggaagctgtatttgcacagtattcat
gccagccaccgccaaacccaagcgctacaacccccgccaccccgaacgcactctgctctaccaa
accgtggccgagcactacgagagctggctggagctggcttgcgagggtcaatttga

Annex 3. Quality data for 16S rRNA gene amplicon sequences

Sequencing quality data for Chapter 2

Sample accession[1]	Alias	Sample Type	Treatment	Sampling time (d) or stage	Original reads	Reads after QC[2]
SAMEA5205738	EJM	Enrichment	NA	P4	179 502	146 565
SAMEA5205737	ECM	Enrichment	NA	P4	37 122	31 235
SAMEA5205741	ESM33	Enrichment	NA	P4	125 897	107 878
SAMEA5205742	ESN37	Enrichment	NA	P4	155 493	120 172
SAMEA5205739	EHM35	Enrichment	NA	P4	143 259	115 653
SAMEA5205740	EHN36	Enrichment	NA	P4	151 874	124 695
SAMEA5205703	JMET1	Soil	NA	P4	150 477	103 222
SAMEA5205702	CMET1	Soil	NA	P4	26 929	17 579
SAMEA5205720	SMET01	Soil	Metaldehyde	0	163 530	122 008
SAMEA5205721	SMET02	Soil	Metaldehyde	16	127 230	98 478
SAMEA5205722	SMET03	Soil	Metaldehyde	32	165 725	128 917
SAMEA5205723	SMET04	Soil	Metaldehyde	64	222 253	175 664
SAMEA5205724	SMET05	Soil	No Metaldehyde	0	120 452	89 753
SAMEA5205725	SMET06	Soil	No Metaldehyde	16	97 343	72 945
SAMEA5205726	SMET07	Soil	No Metaldehyde	32	226 851	178 495
SAMEA5205727	SMET08	Soil	No Metaldehyde	64	177 669	139 056
SAMEA5205728	SNO09	Soil	Metaldehyde	0	229 170	177 590
SAMEA5205729	SNO10	Soil	Metaldehyde	16	171 275	132 956
SAMEA5205730	SNO11	Soil	Metaldehyde	32	217 476	169 814
SAMEA5205731	SNO12	Soil	Metaldehyde	64	301 296	241 169
SAMEA5205732	SNO13	Soil	No Metaldehyde	0	142 457	108 783
SAMEA5205733	SNO14	Soil	No Metaldehyde	16	183 186	144 217
SAMEA5205734	SNO15	Soil	No Metaldehyde	32	183 828	142 526
SAMEA5205735	SNO16	Soil	No Metaldehyde	64	117 056	90 379
SAMEA5205704	HMET17	Soil	Metaldehyde	0	336 213	266 286
SAMEA5205705	HMET18	Soil	Metaldehyde	16	146 498	114 922
SAMEA5205706	HMET19	Soil	Metaldehyde	32	112 576	87 551
SAMEA5205707	HMET20	Soil	Metaldehyde	64	117 637	92 318
SAMEA5205708	HMET21	Soil	No Metaldehyde	0	112 892	84 332
SAMEA5205709	HMET22	Soil	No Metaldehyde	16	75 787	56 749
SAMEA5205710	HMET23	Soil	No Metaldehyde	32	126 672	98 517
SAMEA5205711	HMET24	Soil	No Metaldehyde	64	145 004	115 223
SAMEA5205712	HNO25	Soil	Metaldehyde	0	370 221	288 031
SAMEA5205713	HNO26	Soil	Metaldehyde	16	172 478	138 317
SAMEA5205714	HNO27	Soil	Metaldehyde	32	288 284	231 116
SAMEA5205715	HNO28	Soil	Metaldehyde	64	193 485	95 459
SAMEA5205716	HNO29	Soil	No Metaldehyde	0	154 185	116 184
SAMEA5205717	HNO30	Soil	No Metaldehyde	16	302 923	245 194
SAMEA5205718	HNO31	Soil	No Metaldehyde	32	95 459	75 440
SAMEA5205719	HNO32	Soil	No Metaldehyde	64	327 198	264 776

[1] Raw reads for all samples were uploaded to the European Nucleotide Archive under study PRJEB30540

[2] Quality Control: quality filtering, denoising and chimera removal

Sample accession[1]	Alias	Sample Type	Treatment	Sampling time (d)	Original reads	Reads after QC[2]
SAMEA7370266	RSMET-41	Soil	Metaldehyde	0	227379	156210
SAMEA7370267	RSMET-42	Soil	Metaldehyde	0	192587	132930
SAMEA7370266	RSMET-43	Soil	Metaldehyde	0	213463	144090
SAMEA7370267	RSMET-44	Soil	Metaldehyde	64	291904	212456
SAMEA7370268	RSMET-45	Soil	Metaldehyde	64	270247	197123
SAMEA7370269	RSMET-46	Soil	Metaldehyde	64	195230	143901
SAMEA7370270	RSMET-47	Soil	No Metaldehyde	64	259851	190194
SAMEA7370271	RSMET-48	Soil	No Metaldehyde	64	213881	156894
SAMEA7370272	RSMET-49	Soil	No Metaldehyde	64	223244	162607
SAMEA7370273	RSMET-50	Soil	Metaldehyde	96	219883	161723
SAMEA7370274	RSMET-51	Soil	Metaldehyde	96	416991	309659
SAMEA7370275	RSMET-52	Soil	Metaldehyde	96	198668	146504
SAMEA7370276	RSMET-53	Soil	No Metaldehyde	96	177475	128759
SAMEA7370277	RSMET-54	Soil	No Metaldehyde	96	197137	145711
SAMEA7370278	RSMET-55	Soil	No Metaldehyde	96	225033	161247
SAMEA7370281	RSNO-41	Soil	Metaldehyde	0	207388	147336
SAMEA7370282	RSNO-42	Soil	Metaldehyde	0	210949	152392
SAMEA7370283	RSNO-43	Soil	Metaldehyde	0	186130	136531
SAMEA7370284	RSNO-44	Soil	Metaldehyde	64	224208	168550
SAMEA7370285	RSNO-45	Soil	Metaldehyde	64	189952	141994
SAMEA7370286	RSNO-46	Soil	Metaldehyde	64	194275	144419
SAMEA7370287	RSNO-47	Soil	No Metaldehyde	64	193688	137350
SAMEA7370288	RSNO-48	Soil	No Metaldehyde	64	253954	184510
SAMEA7370289	RSNO-49	Soil	No Metaldehyde	64	192657	137686
SAMEA7370290	RSNO-50	Soil	Metaldehyde	96	219230	161113
SAMEA7370291	RSNO-51	Soil	Metaldehyde	96	92521	64225
SAMEA7370292	RSNO-52	Soil	Metaldehyde	96	140401	101208
SAMEA7370293	RSNO-56	Soil	No Metaldehyde	96	237139	183332
SAMEA7370294	RSNO-57	Soil	No Metaldehyde	96	215116	162047
SAMEA7370295	RSNO-58	Soil	No Metaldehyde	96	187316	139261

[1] Raw reads for all samples were uploaded to the European Nucleotide Archive under study PRJEB30540

[2] Quality Control: quality filtering, denoising and chimera removal

Sequencing quality data for Chapter 3

Sample accession[1]	Alias	Soil	Treatment	Sampling time (d)	Original reads	Reads after QC[2]
SAMEA7368590	E4	PA	Metaldehyde	48	208718	135185
SAMEA7368587	E9	PA	No Metaldehyde	48	266649	171014
SAMEA7368588	E11	STA	No Metaldehyde	0	365993	238447
SAMEA7368589	E12	STA	Metaldehyde	4	244125	172308
SAMEA7368591	E17	STA	No Metaldehyde	4	220588	153442
SAMEA7368592	E24	ING	Metaldehyde	48	238930	157170
SAMEA7368593	E29	ING	No Metaldehyde	48	216520	134171
SAMEA7368594	E35	LE	Metaldehyde	64	294815	195496
SAMEA7368595	E40	LE	No Metaldehyde	64	175212	98638

[1] Raw reads for all samples were uploaded to the European Nucleotide Archive under study PRJEB40574
[2] Quality Control: quality filtering, denoising and chimera removal

Sequencing quality data for Chapter 4

Sample accession[1]	Alias	Sample Type	Treatment	Sampling time (d)	Original reads	Reads after QC[2]
SAMEA7370390	SF1A	Sand Filter	No Metaldehyde	1	277365	105595
SAMEA7370391	SF1B	Sand Filter	No Metaldehyde	16	269167	118106
SAMEA7370392	SF1C	Sand Filter	No Metaldehyde	18	275687	121329
SAMEA7370393	SF1D	Sand Filter	No Metaldehyde	23	289785	115987
SAMEA7370394	SF1E	Sand Filter	No Metaldehyde	30	260102	104901
SAMEA7370395	SF1F	Sand Filter	No Metaldehyde	37	175852	68522
SAMEA7370396	SF1G	Sand Filter	No Metaldehyde	42	296811	114049
SAMEA7370397	SF1H	Sand Filter	No Metaldehyde	44	151194	61312
SAMEA7370398	SF1I	Sand Filter	No Metaldehyde	51	158385	63872
SAMEA7370399	SF1J	Sand Filter	No Metaldehyde	56	227660	88644
SAMEA7370400	SF1K	Sand Filter	No Metaldehyde	58	210510	86987
SAMEA7370401	SF1L	Sand Filter	No Metaldehyde	65	363532	148125
SAMEA7370402	SF1M	Sand Filter	No Metaldehyde	72	373628	145823
SAMEA7370429	SF2A	Sand Filter	No Metaldehyde	1	155823	61653
SAMEA7370430	SF2B	Sand Filter	No Metaldehyde	16	234670	145410
SAMEA7370431	SF2C	Sand Filter	No Metaldehyde	18	299934	177252
SAMEA7370432	SF2D	Sand Filter	No Metaldehyde	23	368704	150083
SAMEA7370433	SF2E	Sand Filter	No Metaldehyde	30	333547	135828
SAMEA7370434	SF2F	Sand Filter	No Metaldehyde	37	295752	126419
SAMEA7370435	SF2G	Sand Filter	No Metaldehyde	42	212016	129990
SAMEA7370436	SF2H	Sand Filter	No Metaldehyde	44	299789	184828
SAMEA7370437	SF2I	Sand Filter	No Metaldehyde	51	323267	154483
SAMEA7370438	SF2J	Sand Filter	No Metaldehyde	56	353724	137366
SAMEA7370439	SF2K	Sand Filter	No Metaldehyde	58	320911	121510
SAMEA7370440	SF2L	Sand Filter	No Metaldehyde	65	351003	145047
SAMEA7370441	SF2M	Sand Filter	No Metaldehyde	72	314773	131776
SAMEA7370485	SF3A	Sand Filter	Metaldehyde	1	405157	222068
SAMEA7370486	SF3B	Sand Filter	Metaldehyde	16	291201	201809
SAMEA7370487	SF3C	Sand Filter	Metaldehyde	18	346595	221635
SAMEA7370488	SF3D	Sand Filter	Metaldehyde	23	253015	101638
SAMEA7370489	SF3E	Sand Filter	Metaldehyde	30	364621	191926
SAMEA7370490	SF3F	Sand Filter	Metaldehyde	37	376754	208405
SAMEA7370491	SF3G	Sand Filter	Metaldehyde	42	350725	280779
SAMEA7370492	SF3H	Sand Filter	Metaldehyde	44	365516	289236
SAMEA7370493	SF3I	Sand Filter	Metaldehyde	51	420642	260275
SAMEA7370494	SF3J	Sand Filter	Metaldehyde	56	395477	337724
SAMEA7370495	SF3K	Sand Filter	Metaldehyde	58	373910	316867
SAMEA7370496	SF3L	Sand Filter	Metaldehyde	65	344948	217341
SAMEA7370497	SF3M	Sand Filter	Metaldehyde	72	267654	221242
SAMEA7370498	SF4A	Sand Filter	Metaldehyde	1	250494	122626
SAMEA7370499	SF4B	Sand Filter	Metaldehyde	16	453571	313420
SAMEA7370500	SF4C	Sand Filter	Metaldehyde	18	309296	196736
SAMEA7370501	SF4D	Sand Filter	Metaldehyde	23	460255	249125
SAMEA7370502	SF4E	Sand Filter	Metaldehyde	30	366412	188227
SAMEA7370503	SF4F	Sand Filter	Metaldehyde	37	253438	132814

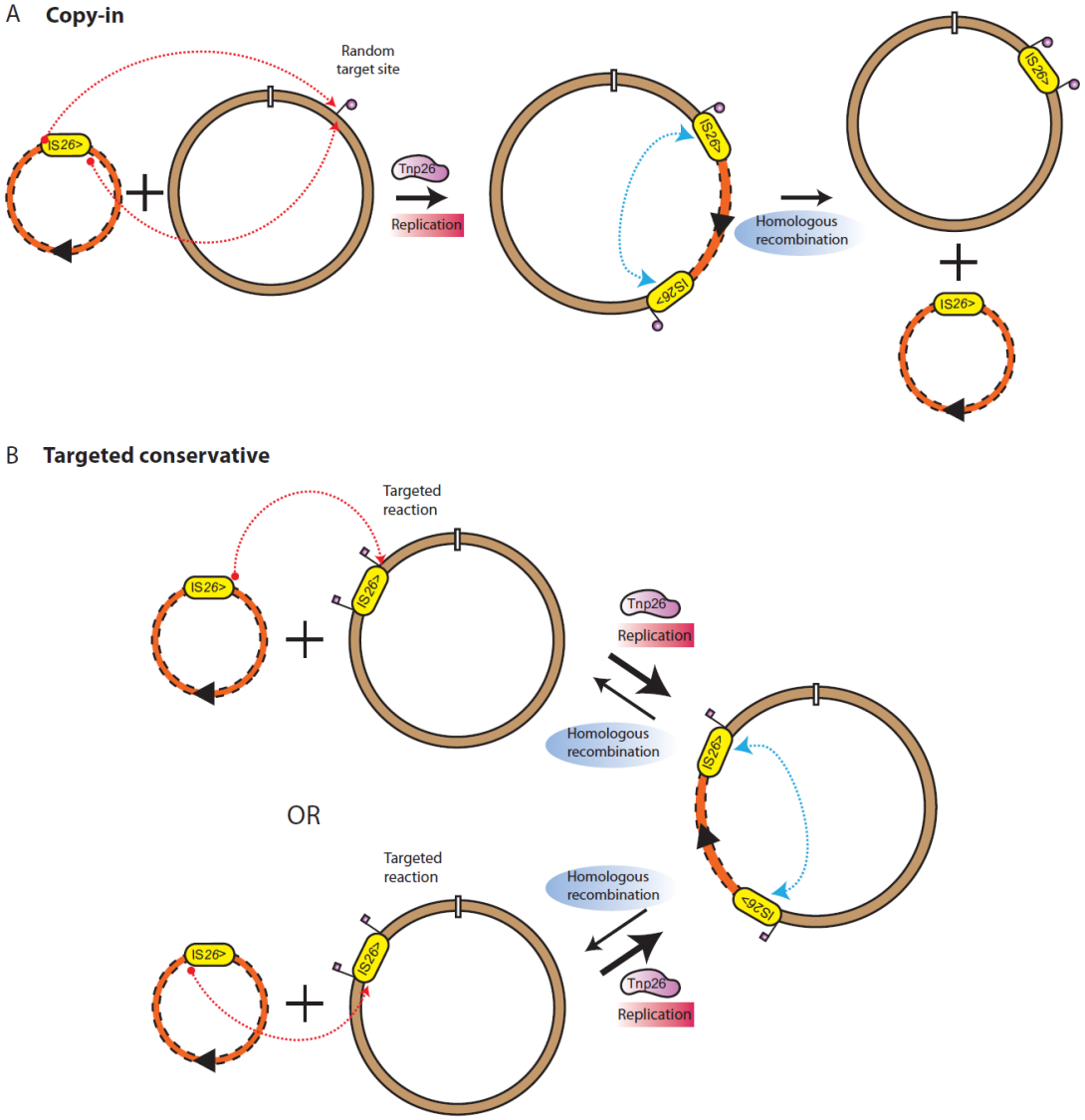
SAMEA7370504	SF4G	Sand Filter	Metaldehyde	42	578597	478209
SAMEA7370505	SF4H	Sand Filter	Metaldehyde	44	255717	203126
SAMEA7370506	SF4I	Sand Filter	Metaldehyde	51	270322	190315
SAMEA7370507	SF4J	Sand Filter	Metaldehyde	56	302852	173970
SAMEA7370508	SF4K	Sand Filter	Metaldehyde	58	241017	155290
SAMEA7370509	SF4L	Sand Filter	Metaldehyde	65	344416	199682
SAMEA7370510	SF4M	Sand Filter	Metaldehyde	72	242037	152656
SAMEA7370511	SF5A	Sand Filter	Metaldehyde	1	367163	187539
SAMEA7370512	SF5B	Sand Filter	Metaldehyde	16	263652	131933
SAMEA7370513	SF5C	Sand Filter	Metaldehyde	18	392321	197340
SAMEA7370514	SF5D	Sand Filter	Metaldehyde	23	116999	20779
SAMEA7370515	SF5E	Sand Filter	Metaldehyde	30	323901	156370
SAMEA7370516	SF5F	Sand Filter	Metaldehyde	37	359140	179208
SAMEA7370517	SF5G	Sand Filter	Metaldehyde	42	342033	157858
SAMEA7370518	SF5H	Sand Filter	Metaldehyde	44	253522	121337
SAMEA7370519	SF5I	Sand Filter	Metaldehyde	51	174902	86001
SAMEA7370520	SF5J	Sand Filter	Metaldehyde	56	359728	179080
SAMEA7370521	SF5K	Sand Filter	Metaldehyde	58	332018	164469
SAMEA7370522	SF5L	Sand Filter	Metaldehyde	65	286768	142095
SAMEA7370523	SF5M	Sand Filter	Metaldehyde	72	247892	122217
SAMEA7370524	SF6A	Sand Filter	Metaldehyde	1	247799	167076
SAMEA7370525	SF6B	Sand Filter	Metaldehyde	16	288952	194173
SAMEA7370526	SF6C	Sand Filter	Metaldehyde	18	266259	178794
SAMEA7370527	SF6D	Sand Filter	Metaldehyde	23	330359	230576
SAMEA7370528	SF6E	Sand Filter	Metaldehyde	30	324147	234176
SAMEA7370529	SF6F	Sand Filter	Metaldehyde	37	165310	116829
SAMEA7370530	SF6G	Sand Filter	Metaldehyde	42	415272	291311
SAMEA7370531	SF6H	Sand Filter	Metaldehyde	44	174407	124436
SAMEA7370532	SF6I	Sand Filter	Metaldehyde	51	319349	227495
SAMEA7370533	SF6J	Sand Filter	Metaldehyde	56	295219	212905
SAMEA7370534	SF6K	Sand Filter	Metaldehyde	58	315118	228006
SAMEA7370535	SF6L	Sand Filter	Metaldehyde	65	325218	235312
SAMEA7370536	SF6M	Sand Filter	Metaldehyde	72	336781	245713

[1] Raw reads for all samples were uploaded to the European Nucleotide Archive under study PRJEB40595

[2] Quality Control: quality filtering, denoising and chimera removal

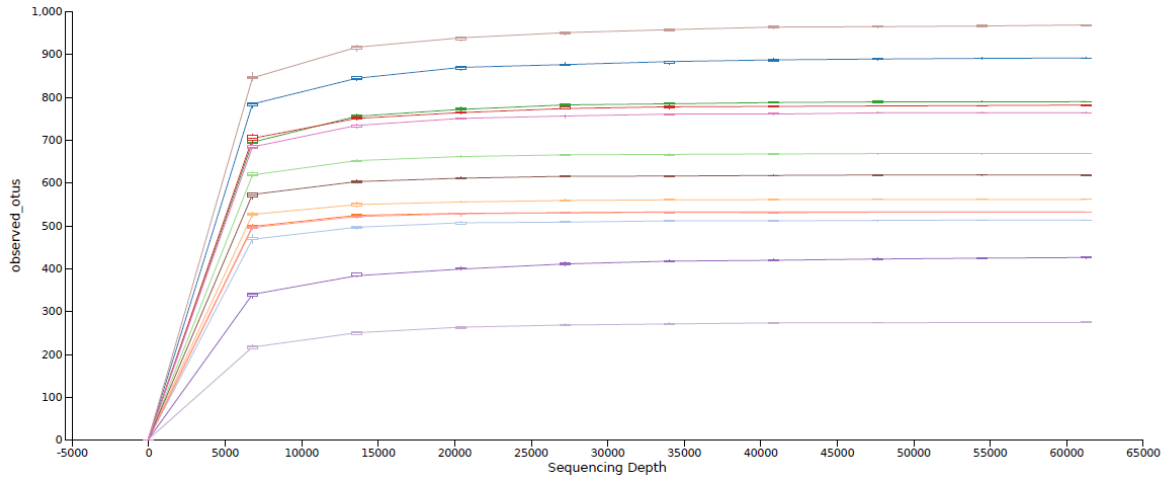
Annex 4. Routes to cointegrate formation by IS26 family members*

*Taken from (Harmer and Hall 2020).

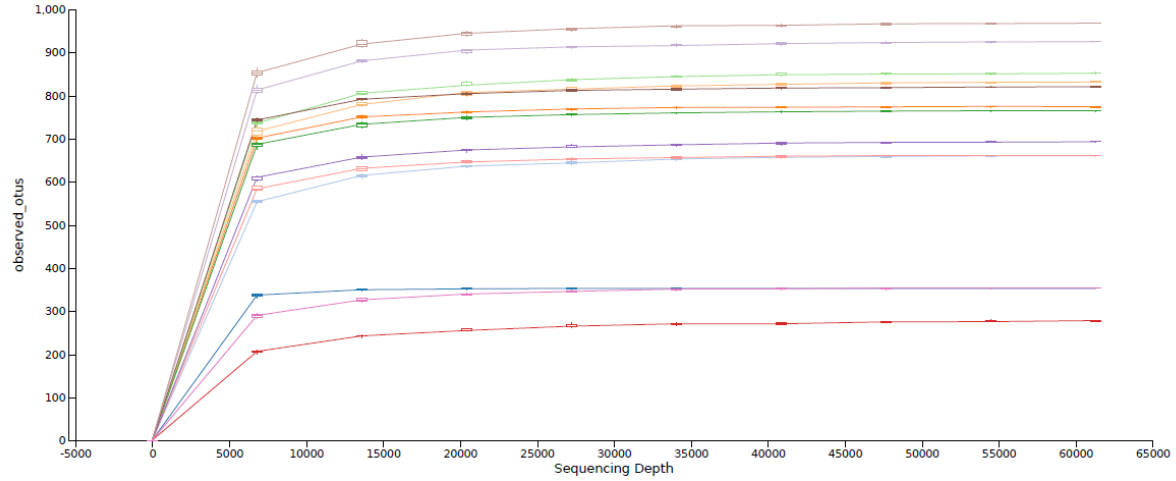


Annex 5. Alpha rarefaction plots for slow sand filters samples showing the number of OTUs as a function of sequencing depth

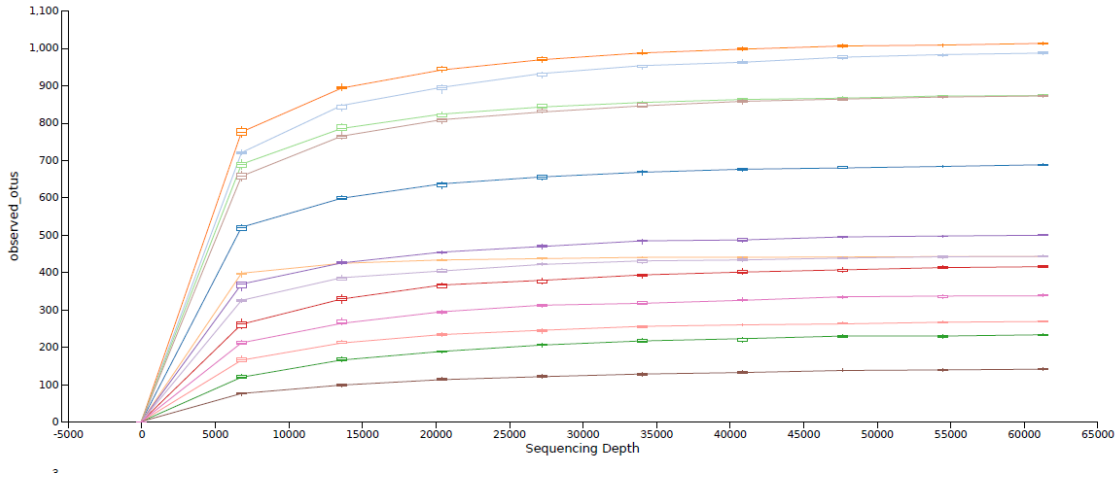
Filter 1



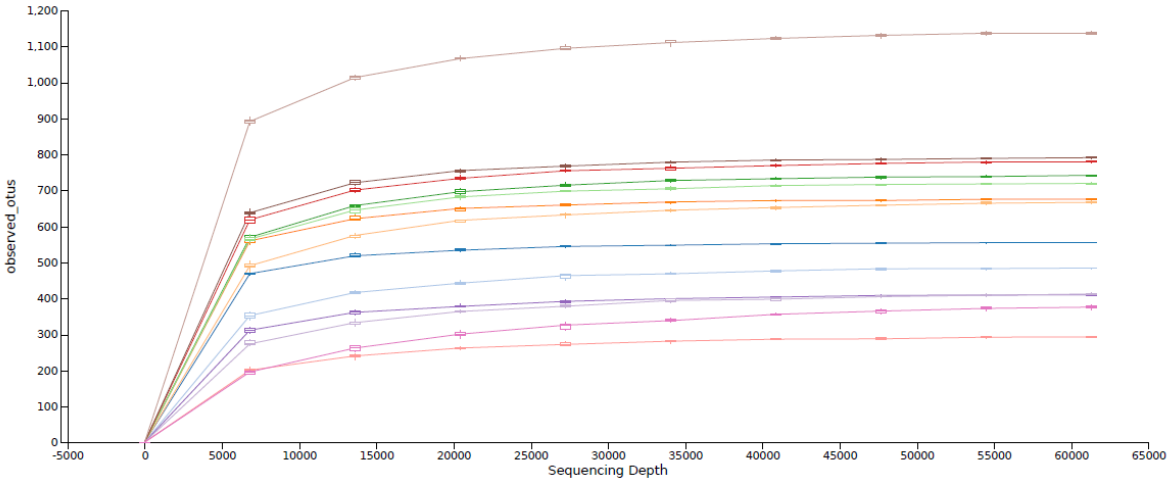
Filter 2



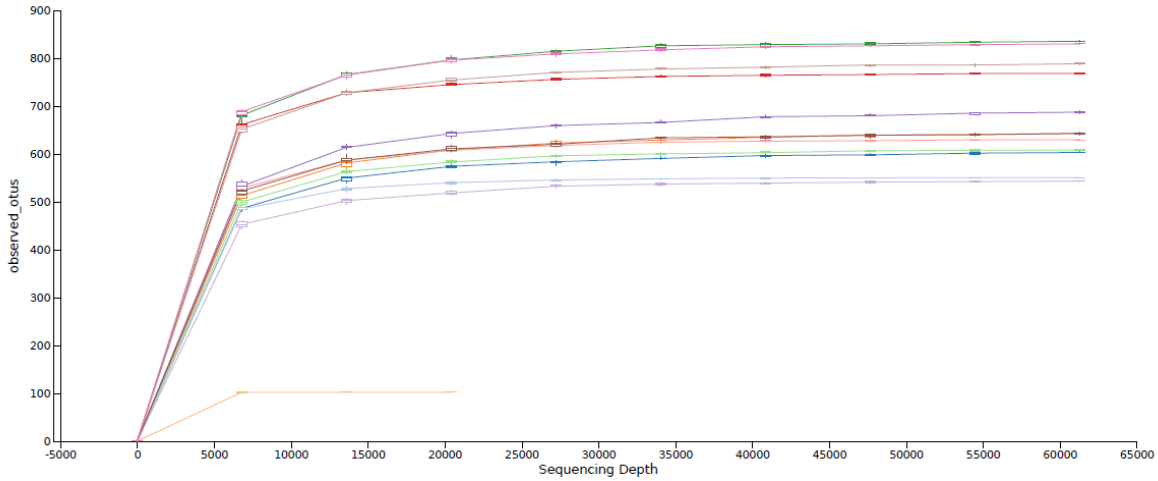
Filter 3



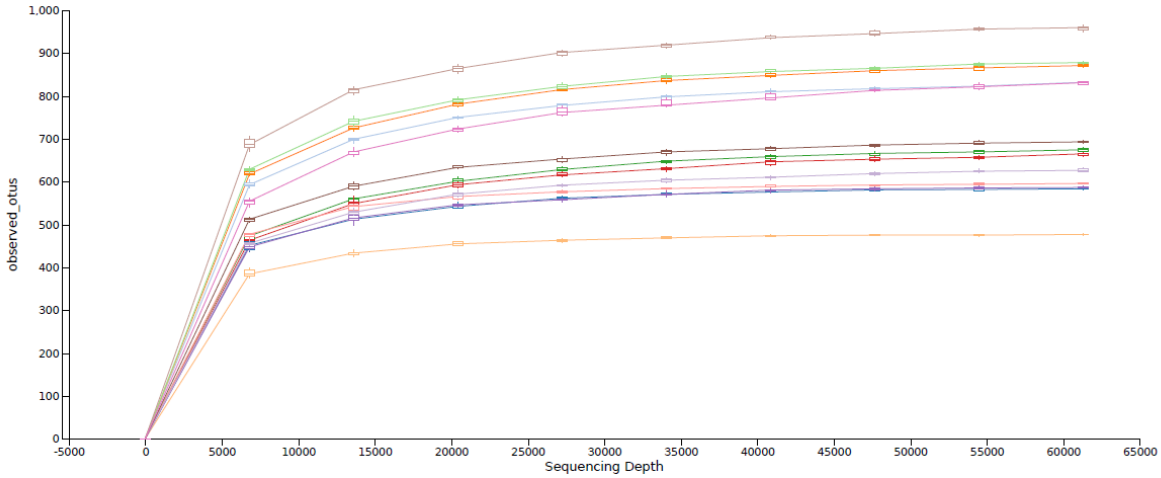
Filter 4



Filter 5



Filter 6



Annex 6. PCO ordination at genus level of bacterial community 16S rRNA in sand filters after restricting the analysis to Proteobacteria only