Isoprene Degradation in the Terrestrial Environment

Gordon Murphy

A thesis submitted for the degree of Doctor of Philosophy

Department of Biological Sciences

University of Essex

2017

Summary

Isoprene production comprises of one third of the total global hydrocarbon release, and influences atmospheric chemistry, including increasing global temperatures and raising ozone concentrations. Although isoprene production, and the chemical attenuation of isoprene have been well studied, microbial degradation of isoprene has had little attention. In this thesis, seasonal isoprene degradation rates of soil were measured, demonstrating an average isoprene consumption rate of 4.77 nmol m^{-2} h^{-1}, which indicates that the effect of soil microbiota acting as a sink for atmospheric isoprene is lower than previously presumed. Neither the season, nor proximity to isoprene-producing trees affected the rate of isoprene degradation in soil, suggesting other isoprene sources in the soil. A series of isoprene enrichment experiments were performed with soil, using the sensitivity of high isoprene concentrations, the evidence of direct isoprene degradation using stable isotope probing, and the relatability to the natural environment using low isoprene concentrations. Different bacteria were enriched at different levels of isoprene, with genera such as *Rhodococcus* being enriched at high concentrations, and *Methylobacterium* being enriched at low concentrations, suggesting specialisation for different isoprene concentrations. Additionally, the clade TM7 was shown to be highly enriched in some isoprene enrichments, although did not incorporate isoprene into its DNA. Changes in bacterial community structure analysis through isoprene enrichment suggested a replicable isoprene-degrading community. An extensive collection of isoprene-degrading isolates was created. Isoprene-degrading genes were investigated at the amplicon and genomic levels. The isoprene-degrading operon displayed variable

chromosomal positioning, appearing in either plasmids or on the main chromosome of closely related species. *IsoA* gene sequence diversity was larger than in previous studies, and the *isoA* gene sequences were not correlated to the *16S rRNA* gene sequences, which could indicate horizontal gene transfer. Isoprene-degrading bacteria were present on a variety of leaves; and communities of bacteria derived from leaves degraded isoprene.

Acknowledgements

I would like to thank my primary supervisor, Dr Terry McGenity, for his advice, support and guidance during this project.

I would also like to thank my other supervisors, Dr Tracy Lawson and Dr Michael Steinke, as well as Professor Colin Murrell, for their help and collaboration, and my board member Dr Alex Dumbrell for his encouragement.

I would also like to thank Farid Benyahia for his help and assistance in the laboratory.

A special thanks goes to my family and friends for their continued support.

Lastly, thank you NERC for making all this possible.

Table of Contents

Chapter 1. Literature review **1**

1.1 Isoprene production **2**

1.2 Effects of isoprene **12**

1.3 Fate of isoprene **15**

1.3.1 Chemical removal **16**

1.3.2 Biological removal **19**

1.4 Isoprene-degrading bacteria **24**

1.5 Isoprene degradation in the terrestrial environment **28**

1.5.1 Aims of the project **30**

Chapter 2. Soil as a sink for isoprene: the effects of season, location, isoprene concentration, and identification of active isoprene-degrading bacteria. **35**

2.1 Abstract **36**

2.2 Introduction **37**

2.3 Materials and Methods **45**

2.3.1 In situ sampling **45**

2.3.2 Ex situ sampling **48**

2.3.2.1 Sequential enrichments with isoprene **48**

2.3.2.2 DNA Stable Isotope Probing of soil with isoprene. **56**

2.3.2.3 Low concentration experiments **58**

2.3.2.4 Amalgamated analysis **58**

2.4 Results: **60**

2.4.1 The rate of isoprene degradation by microbes within soils is at the lower end of previous estimates, with negligible seasonal effects **60**

2.4.2 Microbes within soil degrade isoprene rapidly, but rates are not significantly affected by proximity to isoprene producing trees. **52**

2.4.3 Sequential isoprene enrichment causes significant changes in bacterial community composition **64**

2.4.4 Isoprene causes demonstrable shifts in bacterial community structure **66**

2.4.5 At realistic concentrations of isoprene exposure, enrichment leads to visible but insignificant shifts of bacterial community structure. **69**

2.4.6 Bacterial OTUs potentially involved in isoprene degradation in soils **71**

2.4.7 OTUs of bacterial groups enriched by isoprene addition **80**

2.5 Discussion **85**

2.5.1 in situ rates of isoprene degradation in soils were near the lower measurements, suggesting overestimation of the size of the global soil sink. **85**

2.5.2 No difference in degradation rates is observed between soils from beneath or remote from isoprene-producing trees. **87**

2.5.3 isoprene-degrading bacteria are numerous, varied, and possibly specialised for different isoprene concentrations **88**

2.6 Conclusion **113**

Chapter 3. Isolation of bacteria associated with isoprene degradation and investigation of the isoprene degradation pathway. **117**

3.1 Abstract **118**

3.2 Introduction **119**

3.3 Materials and Methods **121**

3.3.1 Enrichment and isolation of isoprene-degrading microorganisms **121**

3.3.2 Analysis of the 16S rRNA and isoA gene sequences of isoprene-degrading strains **124**

3.3.3 Carbon source testing **125**

3.3.4 Genome sequencing **125**

3.3.5 Attempts at isolating TM7 **127**

3.4 Results **127**

3.4.1 Numerous types of bacteria are able to grow on isoprene as the sole carbon source. **127**

3.4.2 There is no correlation between isoA gene sequence and 16S rRNA sequence **130**

3.4.3 Isoprene-degrading bacteria have a wide range of other metabolic capabilities **133**

3.4.4 Isoprene-degrading Rhodococcus have similar metabolic capabilities, but different genomic arrangements. **135**

3.4.5 The Rhodococcus AD45 isoprene-degrading operon may be the exception **136**

3.4.6 Isoprene-degrading genes from the strains in this thesis were similar to those from known isoprene degraders. **140**

3.4.7 TM7 may be interacting with Aeromicrobium **142**

3.5 Discussion **145**

3.5.1 No phylogenetic linkage observable between isoA gene sequence and 16S rRNA gene sequence **146**

3.5.2 A different, more diverse operon involved in isoprene degradation exists **148**

3.5.3 TM7 may be interacting with Aeromicrobium **148**

3.6 Conclusion **149**

Chapter 4: Isoprene degradation on the leaf surface **152**

4.1 Abstract: **152**

4.2 Introduction: **153**

4.3 Materials and Methods **158**

4.3.1 Experimental Design: Tobacco communities. **158**

4.3.2 Bioinformatics: Tobacco bacterial communities **159**

4.3.3 Enrichment of leaf associated bacteria with isoprene. **160**

4.3.4 Experiment to determine isoprene-degradation on detached oak leaves **162**

4.3.5 Experiment to test the effect on net isoprene production by leaves after inoculation with isoprene-degrading bacteria **163**

4.4 Results **165**

4.4.1 Tobacco leaf community structure is not significantly driven by isoprene production **165**

4.4.2 Degradation of isoprene by on-leaf microbiota is undetectable **167**

4.4.3 The addition of isoprene-degrading bacteria to tree leaves did not lead to a long term difference in net isoprene production **168**

4.4.4 Leaf bacteria degraded isoprene, and isoprene degrading bacteria are present on a wide variety of leaves **169**

4.5 Discussion **171**

4.6 Conclusion **174**

Final discussion **176**

References **182**

2.SI: Chapter 2 Supplementary Information **213**

3.SI: Chapter 3 Supplementary Information **322**

2.AX: Chapter 2 Appendix **334**

2.AX.1 Development of a concentration device for volatile hydrocarbons. **334**

2.AX.2 Images of soil volatile chambers **342**

Table of Figures

Figure 1.1. The chemical structure of Isoprene. **3**

Figure 1.2 Isoprene cycle by sources and sinks. **4**

Figure 1.3, MEP and Mevalonate pathways to isoprene synthesis. **9**

Figure 1.4 Lifespan of atmospheric compounds. **17**

Figure 1.5 The indirect mechanism by which isoprene causes increased ozone concentrations. **19**

Figure 1.6 The effect of isoprene concentration on isoprene uptake. **23**

Figure 1.7. Proposed *Rhodococcus* isoprene degradation pathway**. 26**

Figure 1.8: 8456bp DNA region of *Rhodococcus* AD45. **27**

Table 1.1: summary of previously isolated isoprene degrading bacteria **28**

Table 2.1 Location, Tree type and basic characteristics of isoprene enrichment soil sample origins. **49**

Figure 2.1, Isoprene degradation rates for soil at Wivenhoe Park, UK. **61**

Figure 2.2, Time for 80% isoprene degradation for microcosms containing 1 g of soil from under Poplar, Willow, Oak, Ash, and Birch trees, and an area with no trees , incubated with 0.1 ml of saturated isoprene headspace. **63**

Figure 2.3, Time for 80% isoprene degradation for microcosms containing 1 g of soil from under the canopy of Poplar, Willow, Oak, Ash, and Birch trees, and an area with no trees, incubated with 1.0 ml of saturated isoprene headspace. **63**

Figure 2.3 Soil enrichment NMDS plot. **65**

Figure 2.4 Detrended Correspondence Analysis of DNA SIP of soil data. **68**

Figure 2.5; Zoomed in section of Figure 2.4 - a Detrended Correspondence Analysis of DNA SIP of isoprene enriched soil. **69**

Figure 2.6 Non metric multidimensional scaling plot of bacterial communities of willow and birch soil incubated with 0, 15 and 150 ppb isoprene. **70**

Figure 2.7 *Rhodococcus* relative abundances in sequential enrichment of soil with isoprene. **72**

Figure 2.8 *Rhodococcus* relative abundances in ^{13}C fractions net of *Rhodococcus* relative abundances in correspondent ^{12}C fractions. **73**

Figure 2.9 *Rhodococcus* relative abundances after incubation with 0, 15, and 150 ppb isoprene. **74**

Figure 2.10 *Saccharibacteria* (aka TM7) relative abundances in sequential enrichment of soil with isoprene. **75**

Figure 2.11 *Saccharibacteria* (aka TM7) relative abundances in ^{13}C fractions net of *Saccharibacteria* relative abundances in correspondent ^{12}C fractions. **76**

Figure 2.12 *Saccharibacteria* (aka TM7) relative abundances after incubation with 0, 15, and 150 ppb isoprene. **77**

Table 2.2 Summary of OTUs showing enrichment in multiple experimental arms, with at least one significant result by sequential enrichment, DNA SIP or low concentration enrichments. **78**

Figure 2.13 Evolutionary relationship between OTUs enriched by isoprene addition, and their nearest type strains. **83**

Figure 3.1, Isoprene degradation pathway of *Rhodococcus* AD45. **120**

Table 3.1 Table categorising isolates. **128**

Figure 3.2 *16S rRNA* gene sequence maximum composite likelihood tree for 16s rRNA. **129**

Figure 3.3 Phylogenetic tree based on *isoA* nucleotide sequences. **131**

Figure 3.4 Pairwise distances for *isoA* gene sequences against *16S rRNA* gene sequences. **133**

Table 3.2 Significant increases in respiration for isolates with different carbon source additions. **134**

Table 3.3 Key characteristics of genome sequenced strains **136**

Figure 3.5 Genome representation of bl28ba cluster most similar to the *Rhodococcus* AD45 isoprene degrading cluster. **137**

Figure 3.6 Genome representation of bl28a cluster most similar to the *Rhodococcus* AD45 isoprene degrading cluster. **138**

Figure 3.7 Genome representation of 1(2b)b cluster most similar to the *Rhodococcus* AD45 isoprene degrading cluster. **139**

Figure 3.8 Genome representation of *Rhodococcus* AD45 cluster. **140**

Figure 3.9 Predicted metabolic functions of TM7. **143**

Figure 3.10 The proteins involved at the terminal and initial steps of TM7 metabolism. **145**

Figure 4.1 Initial isoprene-producing tobacco experimental layout. **159**

Table 4.1 Tree location for Wivenhoe Park samples. **162**

Figure 4.2 Photograph of leaf chambers, with leaves, without foil coverings. **163**

Figure 4.3 Bacterial leaf communities for isoprene-producing and non-producing tobacco after six weeks plant growth. **166**

Figure 4.4 Changes in isoprene concentration on detached oak leaves. **167**

Figure 4.5 Net production of isoprene with treated and untreated leaves following a six week settling period. **168**

Table 4.2 Isoprene enrichments of cells detached from tree leaves showing a decrease in isoprene after 15 days. **169**

Figure 4.6 Isoprene concentration over time for isoprene enrichments containing microbiota removed from isoprene-producing tobacco. **170**

List of abbreviations

ATP: Adenosine triphosphate

BVOC: Biogenic Volatile Organic Compounds

CTAB: Cetyl trimethylammonium bromide

DCA: Detrended correspondence analyses

DFP: 1,1,1,2,3,4,4,5,5,5-Decafluoropentane

DMAPP: Dimethylallyl pyrophosphate

GC-FID: Gas Chromatograph with Flame Ionisation Detector

GMBA: 2-glutathionyl-2-methyl-3-butenoic acid

IPP: Isopentenyl pyrophosphate

isoMO: Isoprene monooxygenase

MEP: Methyl D-erythritol 4-phosphate

NADPH: Nicotinamide adenine dinucleotide phosphate

NMDS: Non metric multidimensional scaling

NMHC: Non-Methane Hydrocarbons

NOx: Mono-nitrogen oxides

OTU: Operational taxonomic unit

PEG: Polyethylene glycol

PTFE: Polytetrafluoroethylene

QQ: Quantile quantile

ROS: Reactive oxygen species

SD: Standard deviation

SE: Standard Error

SIP: Stable isotope probing

Tg: Teragram

UV: Ultra violet

VD: Vapour density

VOCs: Volatile Organic Compounds

Chapter 1. Literature review

Atmospheric pollution is currently one of the major threats to human health and the environment (WHO, 2014). Due to this, it is important to gain a thorough understanding of the cycling of atmospheric pollutants. This research focusses on the degradation of one of the significant atmospheric pollutants, isoprene, which, compared to other similar compounds, has been markedly understudied.

Our meagre knowledge of the biological degradation of isoprene creates difficulties in predicting future effects of isoprene production on the climate, and similarly the effects of climate on isoprene production. This research adds clarity to this situation, improving our understanding of isoprene flux, and identifying and characterising the key isoprene-degrading bacteria.

This chapter reviews the main processes involved in: (i) The production of isoprene, (ii) the effects of isoprene on the climate and human health, (iii) the attenuation of isoprene, and (iv) the interaction between isoprene and microorganisms, as well as how this research expands our knowledge and abilities in the area.

1.1 Isoprene production

Plants emit more hydrocarbon into the atmosphere than human activity does directly, and significantly more when temperatures are elevated (Purves et al., 2004; Sharkey et al., 2008). Around 1150 Tg of these hydrocarbons are Biogenic Volatile Organic Compounds (BVOC) (Atkinson & Arey, 2003; Guenther et al., 1995). Non-Methane Hydrocarbons (NMHC), short chain (<10 carbon) volatile carbon based molecules, make up most of these BVOCs (Shaw, 2001), and Isoprene is one of these NMHCs (Rasmussen & Went, 1965), and represents approximately one third of total global hydrocarbon

release. Although there is slight variation in estimates, isoprene production is in the region of 600 million tonnes of carbon per annum, on par with global methane production (Guenther et al., 2006). Terpenes, which are made up of isoprene subunits, are the other major category of NMHC, with annual global production of 480 Tg.

Isoprene is also known as 2-methyl-1,3-butadiene (C_5H_8) and: i) has low water solubility, ii) is non polar, iii) has vapours more dense than air, iv) is volatile with a boiling point of 33°C, v) has a mass of 68.0626 g/mol, vi) is uncharged, and vii) is highly reactive due to its two carbon-carbon double bonds (Figure 1.1) (NCIB, n.d.; Seinfeld & Pandis, 2012). This reactivity means that it has important effects on atmospheric chemistry, and so a thorough understanding of its flux is required to inform climate models (Pacifico et al., 2009).

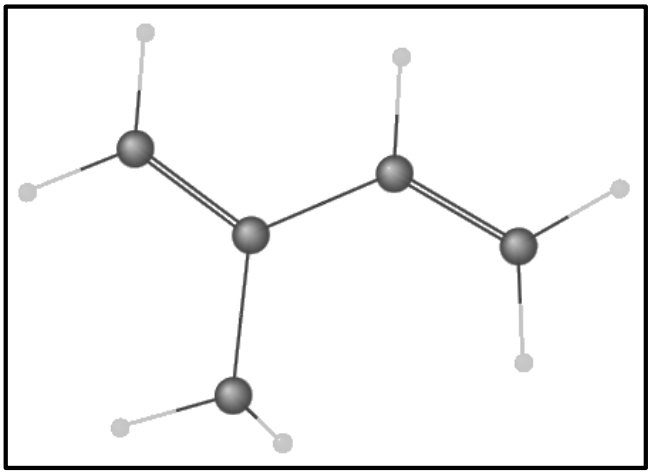


Figure 1.1. The chemical structure of Isoprene. (CH_2=C(CH_2)CH=CH_2). Dark grey circles represent carbon atoms, light grey circles represent hydrogen atoms. Modified from http://pubchem.ncbi.nlm.nih.gov/compound/isoprene.

The vast majority of isoprene is produced by plants. Trees and shrubs are particularly represented, with broadleaf trees accounting for 51% of total production, and shrubs producing 46% (Figure 1.2) (Guenther et al., 2006).

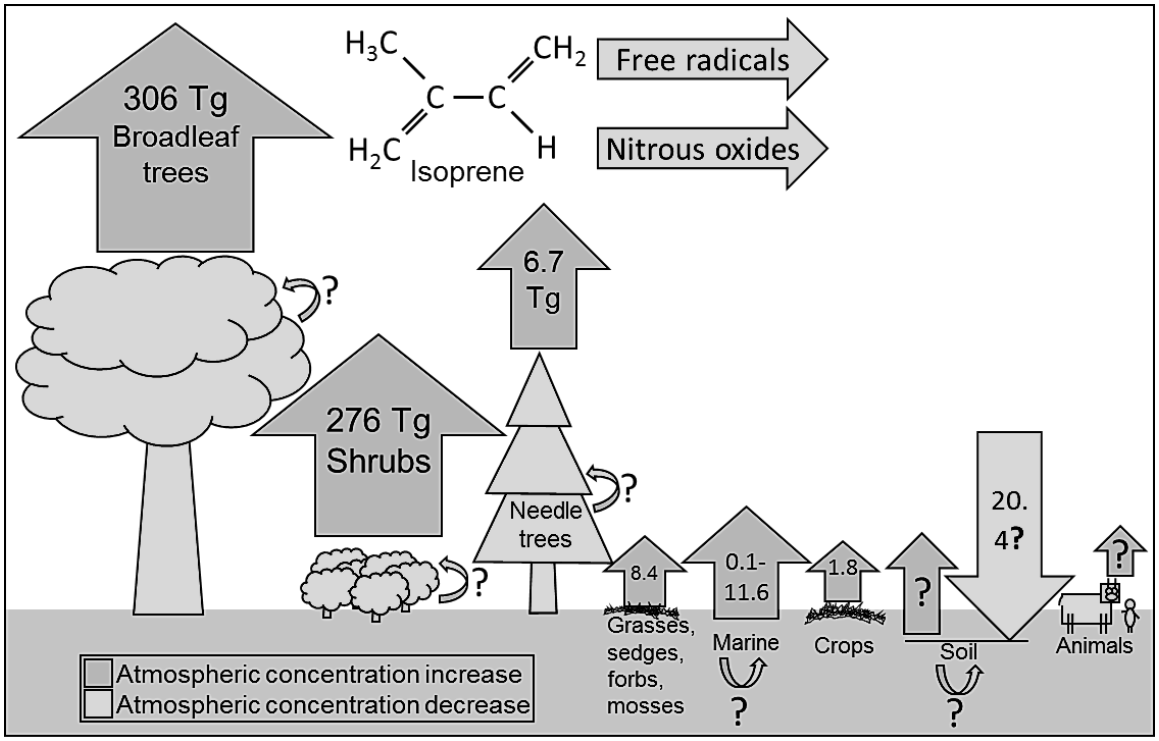


Figure 1.2 Isoprene cycle by sources and sinks. No global data was available for soils, animals or bacteria, although it seems unlikely that these have a significant impact (Wagner et al., 1999; Guenther et al., 2006; Arnold et al., 2009).

This isoprene production is not evenly distributed within each of the sources shown in Fig. 1.2, with a wide range of production rates, and with many species of trees and shrubs not producing isoprene (Sharkey et al., 2008). Taking oaks as an example, analysing the maximum reported production for the 60 oak species recorded in the Hewitt database on VOC emissions, which represents a sample of 10% of oak species (Hewitt, n.d.; Mabberley, 1987); 37% of oak species do not produce isoprene (<0.05 $\mu g^{-1}gdw^{-1}h^{-1}$), and those that do are distributed over the complete range from 0.1 to 151 $\mu g^{-1}gdw^{-1}h^{-1}$ (Average 44.6, SD 43.4). This lack of conformity between taxonomy and trait, together with the fact that species such as Poplar can have multiple isoprene synthase genes, suggest that the ability to synthesise isoprene has evolved separately many

times (Harley et al., 1999; Sharkey & Yeh, 2001), and it has even been postulated that isoprene may originally have been synthesised by the first land plants to better deal with the terrestrial environment (Hanson et al., 1999).

The global distribution of isoprene emitting plants is unequal, with the majority (80%) of isoprene being produced in tropical forests, mid level emissions from temperate forests, and low emissions from boreal forests (Arneth et al., 2011; Pacifico et al., 2009). These emissions are diurnal, peaking around mid-day, mainly due to light dependency (Monson & Fall, 1989), with little night time production (Geron et al., 2012). Isoprene production is also temperature dependant, starting after leaves (the main site of isoprene release) are exposed to around 20°C and increasing until 40-42°C (Rasulov et al., 2010). On a global scale, this leads to seasonal changes in isoprene emissions (Geron et al., 2012; Guenther et al., 2006), and can result in up to ten times more isoprene being produced in the summer than in winter months, likely due to the loss of leaves from broadleaf trees in winter (Rasmussen & Went, 1965).

Plants produce isoprene at a cost, with 0.3 to 5% of carbon fixation going to isoprene production in most isoprene-producing plants – reaching 8% in some, and up to 50% under drought conditions (Baldocchi et al., 1995; Dani et al., 2014; Pegoraro et al., 2005), as well as requiring 20 molecules of ATP and 14 molecules of NADPH per isoprene molecule (Sharkey & Yeh, 2001). This metabolic burden suggests that isoprene production must yield significant advantages for the plants that produce it.

The most likely reason for isoprene production is as a response to transient heat stress (Sharkey & Yeh, 2001). This is supported by: (i) Isoprene addition

restoring heat stress recovery to plants with artificially inhibited isoprene production (Sharkey et al., 2001), (ii) Poplar, engineered to suppress isoprene production, showing poor recovery after heat stress (Behnke et al., 2007), (iii) The correlation between leaf height, and therefore lack of shading, and isoprene production (Sharkey et al., 1996), (iv) The correlation between isoprene-producing species and tropical environments (Guenther et al., 2006), (v) The increased isoprene production when leaves are exposed to light or heat (Sasaki et al., 2005), and (vi) Inhibition of isoprene production reducing heat shock recovery of leaves, and exogenous application of isoprene restoring the ability of plants to recover from heat shock (Sharkey et al., 2001). Isoprene only protects against short term heat stress as demonstrated by the lack of protective effect after long term heat stress (Sharkey et al., 2001; Sharkey et al., 2008).

The temperature of plant leaves can fluctuate rapidly, often to over 10°C above ambient temperatures, potentially more than a hundred times a day (Sharkey & Yeh, 2001). These fluctuations could be severely detrimental to plant health, for example with the temperatures reached being potentially damaging to photosystems by inducing thylakoid membrane leakage, and even dissociation above 45°C (Gounaris et al., 1984; Schrader et al., 2004; Sharkey & Yeh, 2001).

Isoprene reduces the damaging effects of temperature. Due to its hydrophobicity, it can intercalate into the middle of the phospholipid bilayer of the thylakoid membranes, increasing their stability and therefore decreasing the likelihood of leaks or separation (Sharkey et al., 1996; Siwko et al., 2007; Species et al., 1997; Velikova et al., 2011). The intercalation of isoprene into

these membranes causes an increase in stability which is equivalent to that of the leaf being 10°C cooler according to the model of Siwko et al. (2007).

It is thought that the reason that isoprene production is more common in the tropics, is due to limits on transpiration rates, caused by the relative humidity outside the plant. As water loss is one of the other primary mechanisms for plants to reduce their temperatures, isoprene production becomes more important (Sharkey & Yeh, 2001).

Isoprene production acting as a mechanism to protect leaves against transient heat stress has become the commonly accepted theory. However, other factors may be involved, for example: (i) Isoprene helps to quench reactive oxygen species in the leaves, which otherwise oxidise and damage the plant cells; although it is now thought that, as ozone (a ROS) decreases isoprene production, any effect is likely to be co incidental (Loreto et al., 2001; Vickers et al., 2009), (ii) Isoprene could play a role in moderating insect herbivory (Laothawornkitkul et al., 2008; Loivamäki et al., 2008), and (iii) Isoprene production could function as a disposal method for excess Dimethylallyl pyrophosphate (DMAPP) (Logan et al., 2000), although this is unlikely to be significant due to the tight regulation of DMAPP production (Banerjee & Sharkey, 2014; Cordoba et al., 2009).

Isoprene is produced through either the methyl D-erythritol 4-phosphate (MEP) pathway or the mevalonate (non-MEP) pathway. The MEP pathway is also called the non-melavonate pathway. The MEP pathway, which converts Pyruvate and Glyceraldehyde-3-phosphate to DMAPP/IPP, is present in the

plant cytosol, archaea, animals, fungi, and some bacteria. The mevalonate pathway, which begins with Acetyl-CoA, is present in cyanobacteria, bacteria and chloroplasts (Loreto et al., 2004; Sanadze, 2004; Sharkey et al., 1991; Pulido et al., 2012; Schnitzler et al., 2004; Trowbridge et al., 2012). In both pathways, the product IPP can be converted to DMAPP by IPP Isomerase, and DMAPP can be converted into isoprene by Isoprene synthase (Figure 1.3).

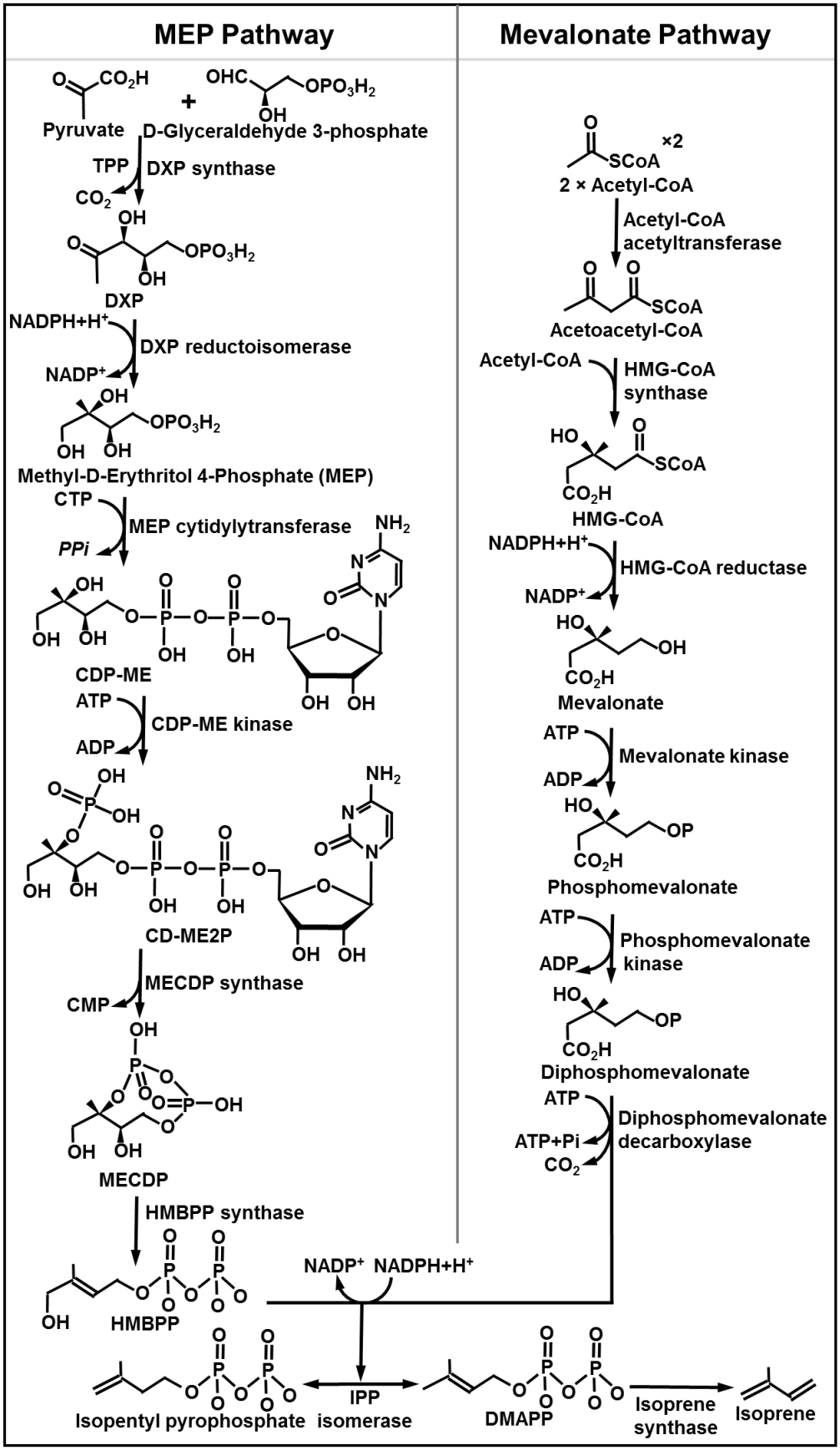


Figure 1.3, MEP and Mevalonate pathways to isoprene synthesis, based on Cordoba et al. (2009); Kuzuyama, 2002; Silver & Fall, 1995)

In plants, and unlike other terpenoids, isoprene is not stored in the leaves (Sharkey & Yeh, 2001), so is produced on demand, which is dependant on the contributing factors (e.g. temperature fluctuations). However, isoprene production levels and isoprene synthase concentrations are not correlated beyond presence/absence, with stress increasing isoprene production, but not isoprene synthase levels in Poplar (Fall & Monson, 1992; Vickers et al., 2010). On the contrary, isoprene synthesis is regulated through enzyme kinetics and substrate supply (Vickers et al., 2009). The high K_m of isoprene synthase suggests that the substrate (DMAPP) concentration is the main regulator (Silver & Fall, 1995; Wiberley et al., 2009). DMAPP production is, at least in the MEP pathway, heavily regulated at the transcription and expression levels (Rosenstiel et al., 2002; Sharkey & Yeh, 2001; Vickers et al., 2010; Wildermuth & Fall, 1996; Wolfertz et al., 2004). The rate of isoprene synthase activity is altered by: (i) light, which generally increases synthesis, (ii) temperature, which generally increases synthesis (to a point), and (iii) CO_2, which increases or decreases isoprene synthase activity dependant on tree species (Arneth et al., 2007; Heald et al., 2009; Logan et al., 2000; Loreto & Sharkey, 1993; Monson & Fall, 1989; Singsaas & Sharkey, 2000; Tingey et al., 1981).

Isoprene is also produced in many bacteria, especially *Bacillus* species (Kuzma et al., 1995). Additionally, many animals produce isoprene, with humans producing 17 mg day^{-1} on average.

Human activity is changing isoprene production. Changes in land use due to increasing demand for biofuel crops, where fast growing high-isoprene-emitting

trees like poplar are grown, or demand for palm oil, where oil palm are grown. Where high isoprene emitting oil palm trees are grown, there is an increase in isoprene emissions. In 2009, Malaysia comprised of 13% oil palm plantations (from 1% in 1974), which increased to approximately 16% of land mass in 2014 at 5.2 million hectares (Armstrong, 2013; Hardacre et al., 2013; Hewitt et al., 2009, MPOB 2014). This human generated increase, combined with the warming climate's effect on isoprene production, leads to models of a likely substantial increase in isoprene emissions in the future (Guenther et al., 2006; Sanderson et al., 2003a). Biofuel crops generating isoprene provide a quandary between growing trees for renewables to reduce carbon, at the expense of producing isoprene; however new isoprene emission free biofuels are being developed (Behnke et al., 2012).

Atmospheric isoprene concentrations vary between the low parts per trillion up to 10 ppb, mainly depending on the tree species and tree coverage in the area, and the air temperature at the time of sampling (Wagnera and Kuttler, 2014). Higher isoprene concentrations are found in tropical rainforests, and around deciduous forests, for example in Amazonia in 1998 isoprene concentrations ranged from 4 to 10 ppb, and in a forest in Tennessee (USA), isoprene concentrations ranged from 3 to 10 ppb (Baldocchi et al., 1995; Kesselmeier et al., 2000). In urban environments in more temperate climates the atmospheric isoprene concentration is in the region of 0.1 ppb, for example in Essen, Germany daytime isoprene concentrations of isoprene are typically between 0.13 and 0.17 ppb, and around 0.01 ppb at night time (Wagnera and Kuttler, 2014). In Shanghai (China), a warmer urban environment near large forests, isoprene concentrations ranged from 1 to 6 ppb, and in Venezuela isoprene

concentrations reached 3 ppb in a number of locations (Donoso et al., 1996). Low isoprene concentrations can be found near marine environments, between 0.001 and 0.01 ppb (Lewis et al., 1997).

1.2 Effects of isoprene

As isoprene is highly reactive, produced in significant quantities and mostly attenuated in the atmosphere, it is regarded in atmospheric chemistry as a very important biogenic hydrocarbon (Seinfeld & Pandis, 2012).

Isoprene reacts in the atmosphere in a multitude of ways that will be expanded on later; however the main routes are through reactions with hydroxyl radicals (OH) and nitrogen oxides (NOx).One of the major problems with isoprene is that its reactivity with Nitrogen Oxides produces tropospheric (low level) ozone. In the upper atmosphere (stratosphere), ozone provides protection from UV light; but in the lower atmosphere (troposphere), ozone is undesirable as it harms human and plant health and is a greenhouse gas (IPCC, 2001; Sanderson et al., 2003b).

The safe limit of ozone published by the World Health Organisation (WHO) is 60 ppb (WHO 2000), a figure that is sometimes exceeded, particularly in eastern USA and China (Sanderson et al., 2003b). This limit for ozone represents the concentration where there is noticeable lung tissue damage and where asthma can become agitated (Sanderson et al., 2003b). At 80 ppb there is a significant decrease in pulmonary function and increase in respiratory problems, including inflammation and changes in responsiveness, leading to an increase in respiratory related hospitalisations (WHO, 2000). To give an

indication of the size of the effect, an increase of 12.5 ppb of current background ozone levels in Europe has been estimated to increase hospitalisation for respiratory conditions by 5%, and a 20% increase could be caused by a 50 ppb increase in ozone. Longer term exposure to ozone levels between 120 and 250 ppb has been shown to cause changes in the epithelium cells in the airways and connective tissues in the lung, often resulting in fibrotic changes (WHO, 2000). In addition to effects on human, ozone can cause significant damage to ecological systems, with a critical level for vegetation of 200 – 1000 ppb for an hour (WHO, 2000). At current estimates, by 2020 in Europe, yields of wheat could be reduced by 3.5% and maize by 1% (Ashworth et al., 2013). Healthcare and crop losses would amount to $7.1 billion, and the detrimental effect on human mortality may make biofuel plantations an ethical issue (Ashworth et al., 2013). In 1999 and 2000, in China, ozone levels present in the Yantze delta valley, a key agricultural area, were estimated to be responsible for a 20-30 % decrease in winter wheat production (Huixiang et al., 2005). Ozone concentrations of 1000 ppb have been shown to kill a large proportion of microbes exposed, and are likely to have effects at lower levels, causing unknown ecological damage (Dyas et al., 1983). To contextualise all this information, in Europe, there is a background level of isoprene of between 20 and 25 ppb, which sometimes increases to 180 ppb for up to an hour, or 145 ppb for up to 12 hours and is mainly anthropogenic (WHO, 2000).

In addition to the current background levels, Europe is aiming to have 10% of crude oil based fuels replaced by biofuels by 2020; however, due to the types of trees grown for biofuel production, including poplar and eucalyptus, an estimated 1365 deaths could be caused by ozone formed from increased

isoprene reacting with NOx (Ashworth et al., 2013). Ozone is already blamed for 22000 premature deaths per year in Europe (Ashworth et al., 2013). Furthermore, models have shown that isoprene production could increase by 148 -187 Tg per annum in a hundred year time span, depending on vegetation changes, leading to a 10-20 ppb increase of ozone over most continents, and 30-50 ppb increase over China, Korea and Japan (Sanderson et al., 2003b).

The reaction of isoprene with free.radicals has an indirect effect of increasing global warming. As isoprene reacts with free radicals they are removed from the atmospheric pool, preventing them from reacting with other compounds, including methane, which is much less reactive than isoprene (Pacifico et al., 2009). To gain a measure of the scope of this issue, an approximate 9% decrease in isoprene concentration, following the eruption of Mt Pinatubo, was calculated to have increased the hydroxyl radical sink for methane by 5 Tg(CH_4) y^{-1} (Telford et al., 2010). As isoprene can lead to increased global warming, and higher temperatures increase isoprene production, it is possible that positive-feedback loops could be occurring (Duane et al., 2002).

Isoprene also causes the formation of secondary organic aerosols, although there is some discussion as to the climatic effects resulting from the aerosols. Hydrocarbons, in particular isoprene and other terpenes, can be oxidised to compounds with a lower volatility and condense to form particulate matter (aerosols). The aerosols can absorb or scatter solar radiation, changing the distribution of energy in the atmosphere and reflecting radiation from the planet, which causes a cooling effect. Secondary organic aerosols can also change cloud precipitation characteristics and can lead to health complaints (Olcese et al., 2007; Remer et al., 2008; Sharkey et al., 2008).

It has been thought that the secondary organic aerosols from isoprene lead to cloud condensation nuclei and have a cooling effect on the earth (Sharkey et al. 2008; Olcese et al. 2007). However, in more recent studies this has been put under scrutiny; plant chambers containing plants which emit isoprene at a low level were spiked with between 2 and 30 ppb isoprene. The concentration and volume of particles greater than 5 nm (nucleating particles) increased after the removal of isoprene from the system, which suggests that isoprene actually inhibits particle formation and therefore cloud condensation (Kiendler-Scharr et al., 2009), with a resultant warming effect. Adding isoprene also reduces particle number and volume concentration - instead, the nucleation rate seems to be linked to the concentration of OH radicals (Kiendler-Scharr et al., 2009). This is backed up by studies done between large forests, which have shown more new particle formation over forests with low isoprene levels than forests with high isoprene levels, and the new particle formation in the high isoprene forests to be lower than that which is normally expected in the atmosphere, despite normal levels of H_2SO_4. It is thought that 22% of SOAs are the direct result of isoprene (Kanawade et al., 2011).

Although aerosols in the form of polyols, which are created by a reaction between isoprene and OH radicals, have been found around Amazonian rainforests, it is thought that they make up 2 million of the 8-40 million tonne estimate for total global biogenic secondary organic aerosols (Claeys et al., 2004).

1.3 Fate of isoprene

Isoprene production is a somewhat hidden global predicament, with mixed and often undesirable effects coupled with rising production. The reactivity of

isoprene means that it is largely attenuated in the atmosphere, however some of it seems to be biologically degraded.

1.3.1 Chemical removal

Isoprene reacts in the atmosphere in a multitude of ways, as described by Fan & Zhang (2004). Isoprene has an atmospheric lifespan averaging one hour which is medium-to-short in atmospheric terms, and much shorter than the atmospheric lifespan of methane which is over ten years (Figure 1.4); leading it to largely have local initial reactions, within around 100 metres of its source, depending on the prevailing conditions.

The lifetime of isoprene is largely sensitive to the concentrations of OH, O_3 and NO_3^-, and when incubated with typical concentrations of OH, O_3, or NO_3^-, isoprene has a lifetime of 1.7, 31 and 0.8 hours respectively (Seinfeld & Pandis, 2012).

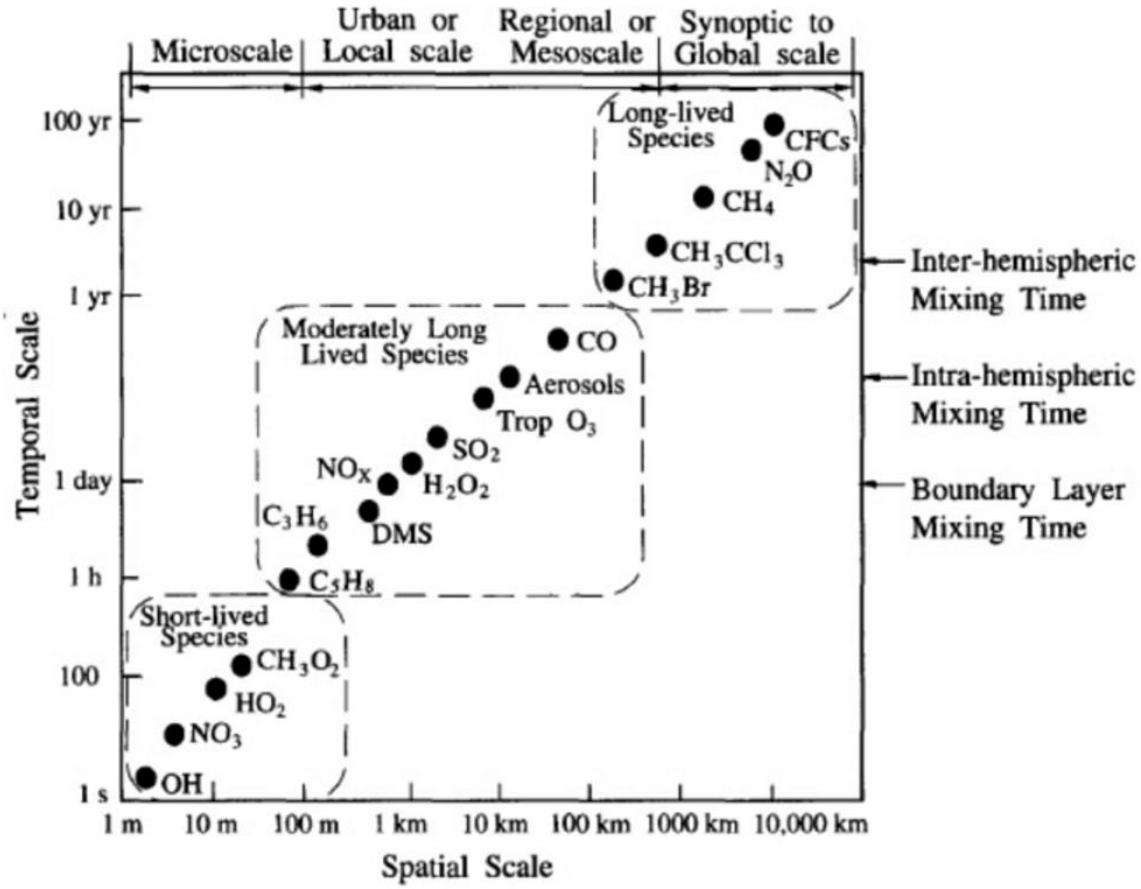


Figure 1.4 Lifespan of atmospheric compounds. Isoprene is indicated by C_5H_8. From Seinfeld & Pandis (2012).

The main attenuation of isoprene is through reactions with hydroxyl (OH) radicals, which are responsible for removal of 85% of the isoprene molecules. Isoprene reacts with hydroxyl radicals in the atmosphere almost entirely by addition reactions, leading to the formation of formaldehyde, including methacrolein and methyl vinyl ketone which are major products alongside carbonyls (Müller et al., 2014; Trapp et al., 2001). As mentioned briefly above, this removal of OH radicals from the atmosphere decreases the concentration of OH radicals, decreasing the rate of the OH-radical-methane reaction.

Much of the rest of the atmospheric sink is through nitrous oxides. Isoprene can react with NO_2 molecules. Through this reaction, outlined in Figure 1.5, isoprene converts NO_2 to NO. The NO is then susceptible to photolysis, leading to the formation of ozone (Fan & Zhang, 2004) (Figure 1.5). That said, it is also possible for isoprene to decrease ozone concentrations in the night time (although isoprene production at night is low) (Fehsenfeld et al., 1992).

Additionally, with nitrous oxides, NO_3 radicals react with isoprene through addition reactions targeting the isoprene carbon-carbon double bond. The resulting addition of O_3 leads to the formation of ozonoids, which decompose to carbonyl and biradical products, generating 1,2 epooxymethyl, fomaldehyde, methacrolein and methyl vinyl ketone (Atkinson & Arey, 2003; Paulson & Seinfeld, 1992).

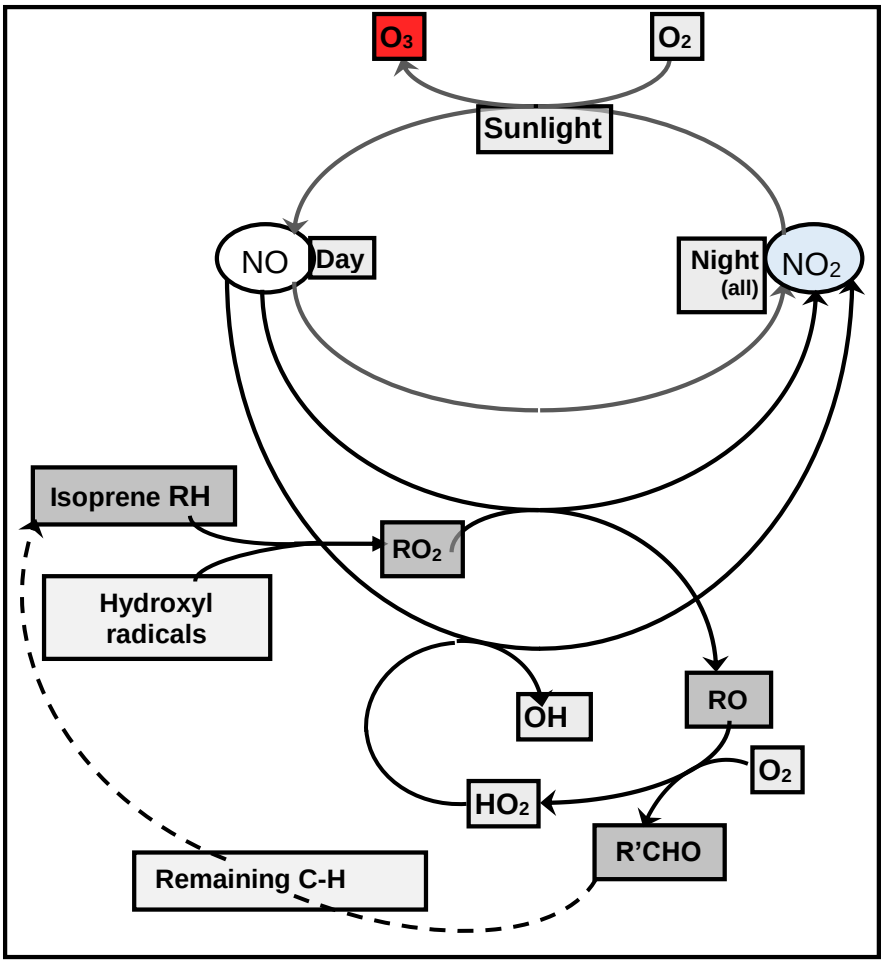


Figure 1.5 The indirect mechanism by which isoprene causes increased ozone concentrations. Loss of one H from an Isoprene molecule, forming two more NO_2 from NO. The net effect is that many molecules of ozone can be formed by one molecule of Isoprene. Image based on information from Sharkey et al. (2008).

1.3.2 Biological removal

Compared to other Volatile Organic Compounds (VOCs), such as methane, little is known about the biological degradation of isoprene.

Notable among the existing studies is Cleveland and Yavitt (1997, 1998), who demonstrated the ability of soils to degrade isoprene, both in situ, and ex situ from a variety of locations, and noted that soil moisture and organic matter were key influencers of high rates of degradation, as well as that isoprene

degradation rates in the soil are highest in the top 3 cm (585 pmol gdw^{-1} day^{-1}), but soil retains the ability to degrade isoprene rapidly down to 9-12 cm (378.81 pmol gdw^{-1} day^{-1}). The key experiment utilised soil chambers spiked with 385 ppb isoprene in a temperate forest. They found that isoprene degradation happened rapidly, often consuming the isoprene to below (their) detection limits (10 ppb) within an hour, and that consumption rates were linear. They also used soil cores with headspace additions of 100 to 700 ppb that were subjected to a variety of temperatures and soil moistures. Isoprene loss in deeper soils (15-18 cm) was much slower than the slowest soils from above that level, with the 12 cm soils being 35% slower than the top layer of soils. Their suggestion was that this could have been due to the decreased organic matter, moisture content, decreased bacterial numbers, and the increased pH associated with depth, or could have been influenced more by isoprene mainly being produced in plants and entering to the soil from the top where the soil contacts the atmosphere.

Cleveland and Yavitt (1998) showed that temperature affected the rates of consumption for the first 7.6 cm of soil, but with no difference between 25 and 30°C. For deeper soils, it had no effect, apart from bringing consumption to around that of the controls at 50°C. The optimum temperature for a degradation rate was indicative of a biological process (c. 30°C). In 24 hours, sterile samples had only consumed 5% of the isoprene, and similarly 5% of the isoprene with no soil was lost, indicating that isoprene consumption was biologically mediated.

Cleveland and Yavitt (1998) also showed that with low O_2 concentrations, consumption rates were significantly diminished (by 80%) and a declining rate

of consumption was observed, providing evidence for the idea that isoprene-degrading micro-organisms are mostly aerobic. They also demonstrated that pretreatment with isoprene enriched the isoprene consuming abilities of soil, evidencing the idea that the consumption of isoprene by soils had a significant microbial contribution.

Soil moisture affected isoprene consumption significantly for upper (0-3 cm) layer soil samples, with 40% and 25% soil moisture content giving the fastest rates, and 100% giving near zero degradation with 6% being the next slowest. This can be explained by the low amount of isoprene that partitions into water making it harder to reach microbes, and the negative effect low moisture content has on microbial activity generally (Cleveland & Yavitt, 1998).

Cleveland and Yavitt (1997) used chambers spiked with 385 ppb isoprene to give an initial estimate of the total global annual isoprene consumption, based on seasonal sampling in the USA, and ex situ sampling in the Americas (USA, Puerto Rico). They also demonstrated rates of 25 pmol g^{-1} h^{-1} at 10 ppb in soils, and through in situ experiments, that higher soil moisture and organic matter lead to increased rates of isoprene degradation, and that the process was biologically mediate.

It was estimated through this work that up to 20.4 million tonnes of isoprene is taken up by the soil, meaning that consumption of isoprene by soil is responsible for around 4% of isoprene emissions (Cleveland & Yavitt 1997). However in this analysis temperate coniferous forest soils were estimated to take up over half of the 20.4 Tg estimate (although with a standard error of 92% of the measurement). This measurement did not take account of the fact that

coniferous trees do not produce much isoprene, instead producing other terpenoids. Cleveland & Yavitt also suggest that isoprene uptake in tropical rainforests was among the lowest of the forest types, despite tropical rainforests being responsible for around 80% of isoprene emissions, suggesting a large scope for error in this value.

It is still unclear, however, whether soil is a net source or sink for isoprene, due to lack of conclusive evidence, as some analyses have shown minimal consumption or low consumption, with night time decreases attributed to chemical reactions and surface deposition (Goldstein et al. 1998; Goldan et al. 1995; Fall & Copley 2000).

Pegoraro (2004 has shown that isoprene consumption of soil is decreased when there is low moisture content, and is not affected by CO_2. Pegoraro's work on CO_2 effects on soil emissions was not designed to determine a rate for isoprene degradation in soils in situ. The (artificially inflated) concentrations of isoprene fluctuate between midday peaks of around 400 ppb, and nearly no isoprene at night, which is far higher than expected in reality, and due to the deliberately, substantially reduced atmospheric reactions in the experimental setup. It is likely that these concentrations would enrich for isoprene-degrading soil microbes. Additionally, the isoprene flux reported for the 430 ppm CO_2, and the lowest tested isoprene concentration (around 100 ppb), yielded an approximate rate of 3600 nmol m^{-2} h^{-1}, although the work was looking at interactions rather than trying to extrapolate rates. Pegoraro's work in this respect does show the insufficiency of using above normal concentrations of isoprene to determine isoprene degradation rates, with rates over five times Cleveland & Yavitt's (1997) value, with both 500 (wet) and 1500 ppb (dry) soil.

Seemingly, extrapolation of Pegoraro's work onto the global soil consumption upper limit suggested by Cleveland and Yavitt, would mean that the soils consume most isoprene produced. However, extrapolating Pegaroro's data back, although it is hard to tell exactly where 15ppb is, on the left of the 50 ppb line added in Figure 1.6 gives an estimate closer to no isoprene uptake, and possibly even little effect of moisture. Confirming this, Gray et al., 2014 shown an isoprene uptake of 2.0 nmol m^{-2} h^{-1} when measuring changes in BVOC concentrations present in the atmosphere at the time of sampling.

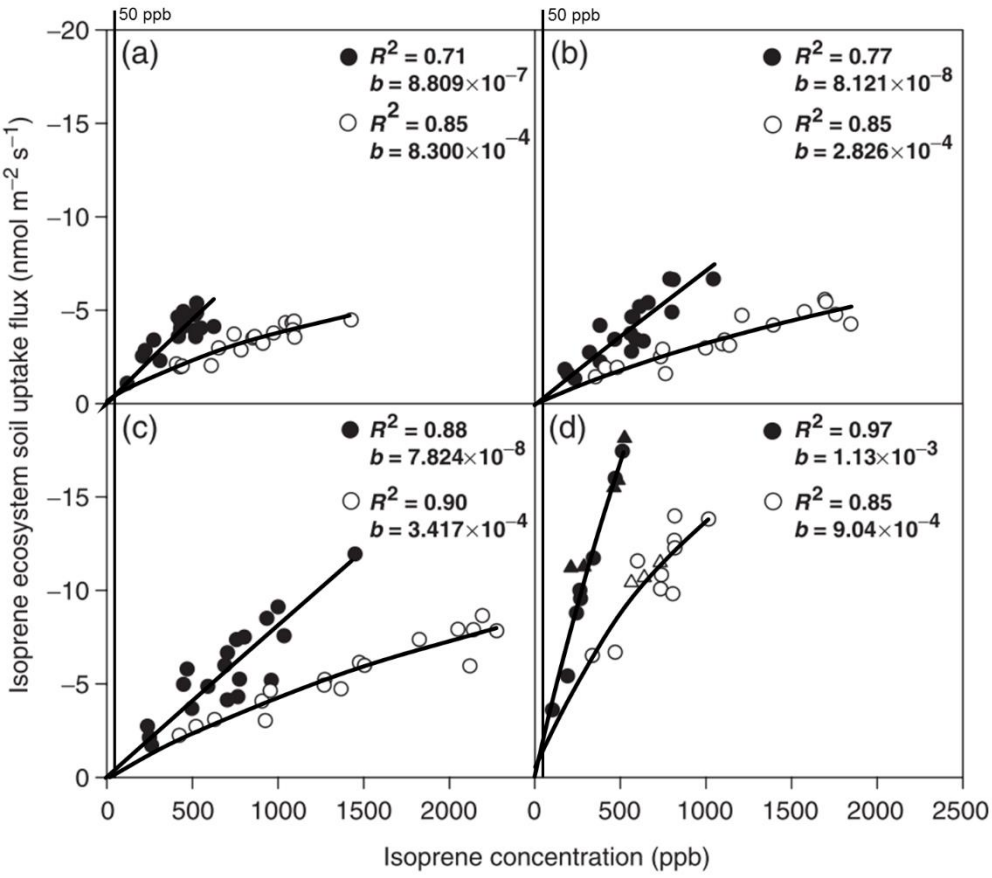


Figure 1.6 The effect of isoprene concentration on isoprene uptake. Edited from Pegoraro et al. (2005) (a: 430 ppm CO_2, b: 800 ppm CO_2, c: 1200 ppm CO_2, d: rainforest mesocosm, black circles = wet, white circles = dry)

1.4 Isoprene-degrading bacteria

The soil layer (< 8 m deep) is the home to 2.6×10^{29} bacterial cells, and is an important part of matter transformation and atmospheric gas exchange (Whitman et al., 1998). One gram of soil can contain up to 10^{10} bacteria, and between 6400 and 38000 species, with a functional and genetic potential which could be more than the cumulative potential of higher organisms (more gene functions), most of which have not been currently cultured in the laboratory, but are metabolically active; for this reason, it is important to gain an understanding of soil microbiology (Torsvik et al. 1996; Curtis et al. 2002).

Due to the high numbers of soil bacteria, and the relevance of soil microbiota in cycling other BVOCs (Ramirez et al., 2009), isoprene-degrading bacteria have so far been assumed to live in the soil ; however as isoprene is largely emitted from leaves, this project has included the investigating of the potential of the leaf surface as a habitat for isoprene-degrading bacteria.

The phyllosphere generally represents a hostile environment for micro-organisms; leaves are subject to rapid fluctuations in temperature, humidity, moisture and levels of ultra-violet light alongside limited nutrients. Diversity and abundance of bacteria depends on the species of plant, the environmental conditions and the age of the leaves (Lindow & Brandl 2003). Despite this, bacteria are present on leaves in concentrations of between 10^6 bacteria cm^3 and 10^8 bacteria cm^3 (10^9 bacteria g^{-1}), alongside lower numbers of archaea and fungi (Beattie & Lindow 1999). The $6.4x10^8$ km^2 of leaf surface globally is host to 10^{26} bacteria (Lindow & Brandl 2003).

Due to the range of conditions, generally versatile bacteria inhabit leaves, with a tendency for pigmented bacteria to thrive better on leaves exposed to more UV light (Jacobs and Sundin, 2001). Some bacteria, such as *Pseudomonas syringae* and *Erwinia* spp., are near ubiquitous in the phyllosphere. Diversity of bacteria on leaves tends to decrease with increased temperature, decreased moisture and leaf age. However, the most important factor in the bacterial community on plants is - carbon sources, with glucose, sucrose and fructose being the most common nutrients originating from the plants. Possibly due to the difficulties in succeeding in colonising leaves, bacteria are found in the highest densities at stomata, joins in the epidermal cell wall, around the veins, and in other plant specific structures and depressions, with significantly more bacteria on the lower surface of the leaf than the upper, where there are more stable conditions (Beattie & Lindow 1999; Lindow & Brandl 2003).

Much of what we know so far about the ability of microbes to degrade isoprene is from culture dependant methods. Following the discovery of *Nocardia* strains that degrade isoprene (van Ginkel et al., 1987), many other species of bacteria have been shown to degrade isoprene in pure cultures, *Alcaligenes denitrifcans* and *Rhodococcus erythropolis* species have been shown to co-degrade isoprene (Ewers et al., 1990). Some *Methanotrophs* and a *Xanthobacter* sp. isolate have also been shown to epoxidise isoprene, but not use isoprene as a sole carbon source (Hou et al. 1981 as cited in Cleveland & Yavitt 2000). *Rhodococcus* AD45 was isolated from fresh water isoprene enrichment cultures and like some of the *Nocardia* strains, can use isoprene as a sole carbon source (Vlieg et al. 2000).

Since then, isoprene degradation has been demonstrated in cultures of *Actinobacteria*: *Gordonia, Leifsonia, Rhodanobacter, Dyadobacter* and *Shinella* like *Alphaproteobacteria*, showing that isoprene-degradation is, phylogenetically speaking, widespread (Acuña Alvarez et al., 2009; Johnston., 2014; El Khawand, 2016), and a recent bioreactor study has shown *Pseudomonas*, *Alcaligenes* and *Klebsiella* to have high efficiency in isoprene degradation (Srivastva et al., 2015).

Table 1.1, Summary of previously isolated isoprene degrading bacteria

Class	Genus
Actinobacteria	*Nocardia*
Actinobacteria	*Rhodococcus*
Betaproteobacteria	*Alcaligenes*
Actinobacteria	*Gordonia*
Actinobacteria	*Mycobacterium*
Actinobacteria	*Leifsonia*
Gammaproteobacteria	*Pseudomonas*
Gammaproteobacteria	*Klebsiella*
Alphaproteobacteria	*Xanthobacter**

* epoxidises isoprene, not used as a carbon source

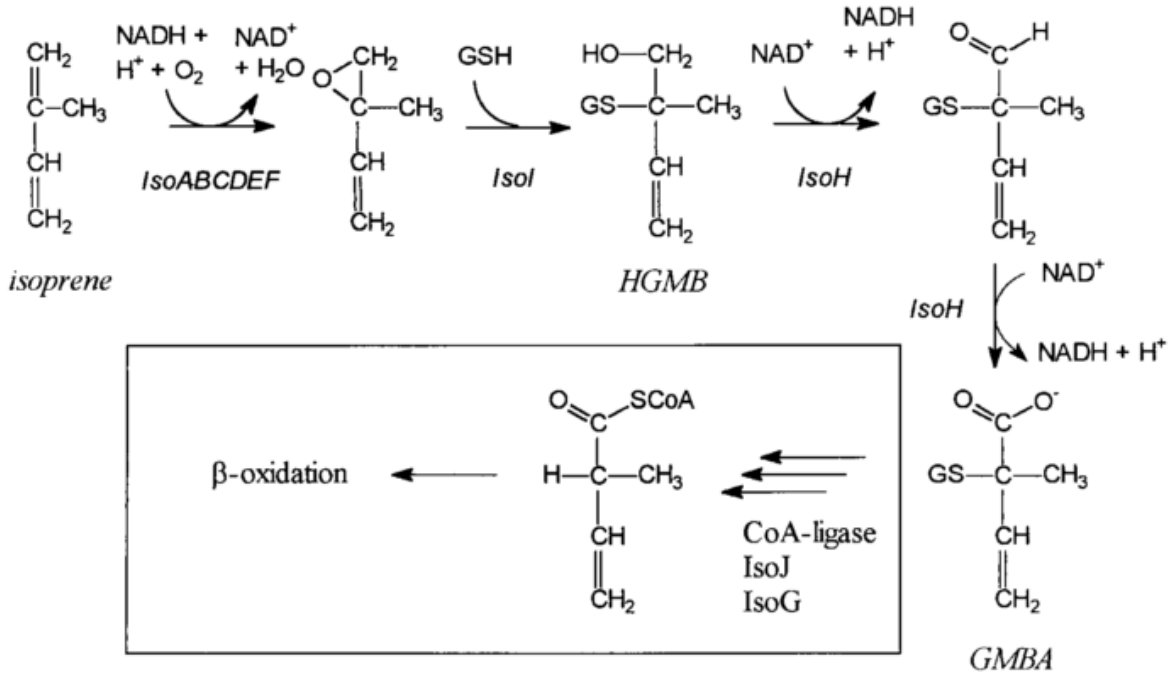


Figure 1.7. Proposed *Rhodococcus* isoprene degradation pathway, with the proposed 2-glutathionyl-2-methyl-3-butenoic acid (GMBA) metabolism pathway in the box. From: Vlieg et al. (2000).

The method by which isoprene is thought to be degraded was elicidated in *Rhodococcus* AD45. It is thought that Isoprene is converted to 1,2-epoxy-2-methyl-3-butene by monooxygenase encoded by *IsoABCDEF*, converted to 1-hydrocy-2-glutathionyl-2-methyl-3-butene (a glutathione conjugate) by a glutathione-S-transferase encoded by *isoI* (prevents epoxide alkylating DNA, proteins etc.) oxidised twice to 2-glutationyl-2-methyl-3-butenioic acid by *isoH* (a unique dehydrogenase, requiring NAD+) through unknown reactions. This may be undertaken by a glutathione conjugate encoded by *isoJ*. i*soG* product could then convert this into 2-methyl-3-butenyl which can be used through central metabolic pathways (Fall & Copley. 2000).

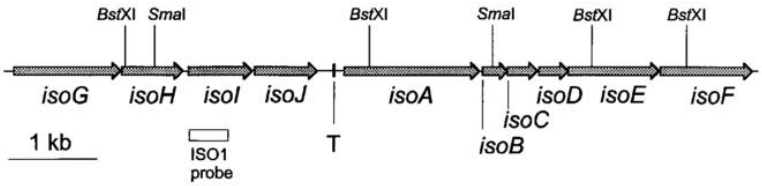


Figure 1.8: 8456bp DNA region of *Rhodococcus* AD45, showing the layout of the *isoA-J* genes, the open reading frame and the rho-independent terminator. From: Van Hylckama Vlieg et al. (2000).

The genes encoding the isoprene-degrading enzyme (IsoMO, a soluble di-iron centre isoprene monooxygenase) in *Rhodococcus* AD45, *isoA-J* (Figure 1.8), are found in a plasmid-encoded isoprene-degrading operon, closely related to a number of the monooxygenases involved in alkene and aromatic oxidation, both in sequence and gene order in clusters (Van Hylckama Vlieg et al., 2000;

Crombie et al., 2015). Expression of the genes encoding IsoMO expression are strongly inducable by one of the products of isoprene degradation, epoxyisoprene, and when induced contribute to 25% of the transcriptome (Crombie et al., 2015).

The isoprene-degrading isolates and the mechanisms and genetics of isoprene degradation are explored further in Chapter 3.

1.5 Isoprene degradation in the terrestrial environment

Isoprene is a reactive, health and climate active gas, and therefore a thorough understanding of isoprene cycling is important - however due to the relative lack of investigation into biological isoprene degradation, there are large areas of knowledge which we lack or which require refining. For example: (1) the isoprene-degradation capacity of soil is largely unknown with current estimates based on an anomaly derived maximum value, (2) the investigations into isoprene-degrading organisms, although useful, have been limited in scope and reproducibility, (3) there has been no previous work (identifiable) on the interaction between isoprene-degrading bacteria and the main source of isoprene, leaves, (4) we know that isoprene-degraders in soil exist, but that they are using the atmosphere as a main source of isoprene has been so far assumed, and (5) the information on the DNA sequences and structures related to isoprene-degradation is deep, but lacks breadth. This work aimed to bring new information, and clarity to these issues, has contributed to building a better picture of the world isoprene degraders live in.

This project (1) refines our knowledge on isoprene degrading rates through in situ chambers with environmentally relevant concentrations of isoprene;, pushing current global estimates of isoprene degradation towards the lower values obtained to date, (2) identifies further isoprene degrading organisms by enrichment, isoprene carbon to DNA incorporation and isolation based methods; expanding our knowledge of isoprene degrading bacterial diversity, creating rich resources for future work, and contextualising previous work, (3) extracts the effect of isoprene-producing tree proximity on soil isoprene degradation rates; indicating that any correlation is unlikely strong enough to have an effect, and that other sources than trees (the main source of atmospheric isoprene) sustain soil isoprene-degrading bacteria, (4) demonstrates the phylogenetic dislinkage between *isoA* genes and *16s rRNA* genes; showing that isoprene degrading ability is either frequently evolved, or (more likely) frequently horizontally transferred, (5) indicates the interactions between plants and isoprene-degraders using transgenic isoprene-producing tobacco; showing either little, or sporadic effects on bacterial community composition, (6) investigates the genomic structure of isoprene degraders in closely related bacteria, showing significant differences in isoprene-degrading cluster placement with genomic integration and plasmid placement, indicating non-permanence in the region, (7) draws out potential bacterial interactions through metagenomic analysis; lending understanding to isoprene-degrading community structuring, and is beginning to explain some unusual enrichments identified in this project. Additionally, this work generated new methods of bacterial isolation, metagenomic investigation; and produced in situ

measurement chambers, a multi-functional gas concentration device, and the associated methodologies.

1.5.1 Aims of the project

This projects aim was to give a wide-ranging overview of isoprene-degradation in the terrestrial environment, in particular to identify and investigate isoprene-degrading bacteria in soils and on leaves, to investigate if isoprene-degrading bacteria were active in the phyllosphere and to add clarity to the effect of soils on the atmospheric isoprene flux. This project was split into three related sections, here as Chapter 2; which looks at isoprene degradation in the soils, and the isoprene-degrading bacteria in the soils, Chapter 3, which investigates isoprene degradation in more depth at the bacterial level, and Chapter 4 which investigates isoprene degradation by bacteria on leaves and the effect of isoprene on the leaf bacterial community.

Chapter 2 was a series of studies designed to elucidate the relevant and active isoprene degrading bacteria in soil and determine the rate of isoprene degradation in soils.

Chapter 2 aimed to:

- Add clarity to the rate of degradation of atmospheric isoprene by soils, and test for seasonal differences. This was investigated by designing and testing chambers and gas concentration systems, deploying the chambers at four points throughout one year, spiking the chambers with an environmentally

relevant concentration of isoprene and measuring the change in concentration over time.

- Test if the mean rate of isoprene degradative capacity was higher in soils from near isoprene producing trees than form soils near non-producing trees. This was investigated by collecting soils from near isoprene-producing trees, and non-producing trees and incubating with isoprene to compare degradation rates.

- Explore the community of isoprene degrading bacteria in the soil through culture-independent methods. This was performed in a three-pronged approach to combine the high sensitivity of high-isoprene concentration serial enrichment of soil, the strong evidence of direct isoprene usage provided by DNA Stable Isotope probing, and the environmental relevance by enrichment of soil with no nutrient enhancement at low isoprene concentrations, all of which was followed by next-generation sequencing and analysis at both the community and OTU level.

- Generate soil derived isoprene-degrading isolates for Chapter 3.

Chapter 3 aimed to investigate isoprene degrading bacteria at the bacteria level. Bacteria, largely derived from Chapters 2 and 4, were isolated and subject to a higher level of analysis of their isoprene-degrading genes, their phylogenetic relationship, ability to degrade other compounds and genomic structures, along with metagenomic analysis of an interesting bacterial community derived from Chapter 2.

Chapter 3 set out to:

- Investigate the diversity of the *isoA* gene, which is key in isoprene degrading bacteria. This was investigated by the isolation of numerous bacteria using isoprene as a sole carbon source through a wide range of isolation techniques. Following isolation, the bacteria's *16S rRNA* and *isoA* regions were sequenced and analysed for their relationship to other bacteria, and for the relationship between the *isoA* and 16S sequence pairs to determine if *isoA* sequences were phylogentically linked.
- Investigate the diversity of the genetic structure of the isoprene degrading operon. This was done through genomic sequencing and analysis of the operon structure and chromosomal position, comparing between bacteria.
- Test the other carbons sources that isoprene-degrading bacteria are capable of using; this was performed through carbon source testing for a number of bacteria, as well as using genomic sequencing and analysis tools for function prediction.
- Investigate the role of TM7 in an isoprene-degrading bacterial community. This was done by metagenomic analysis, function prediction and differential analysis.

Chapter 4 aimed to investigate the isoprene-degrading community present on leaves, and set the foundations of our knowledge on isoprene-degradation occurring in the phyllosphere.
Chapter 4 set out to:
- Test whether bacteria on leaves had any effect on atmospheric isoprene concentrations. This was performed by designing and testing leaf chambers, and incubating leaves in the dark, spiking with isoprene and monitoring the

changes in isoprene concentration, as well as incubation of leaves from isoprene-producing and non-producing trees with isoprene, monitoring and comparing the degradation rates.

- Examine the effect of isoprene production on leaf microbiota. Bacterial cells from the leaves of transgenic tobacco either azygous or heterozygous isoprene production insertions were collected, and had their 16S rDNA regions subjected to next-generation sequencing and analysis.
- Test whether isoprene-degrading bacteria could be used to decrease the net flux of isoprene from the phyllosphere. This was tested by inoculating leaves of isoprene-producing trees with isoprene-degrading bacteria, allowing time for stabilisation and environmental effects, and then testing of isoprene degradation rates.
- Generate leaf-derived isoprene-degrading isolates for Chapter 3.

Chapter 2. Soil as a sink for isoprene: the effects of season, location, isoprene concentration, and identification of active isoprene-degrading bacteria.

2.1 Abstract

One third of global hydrocarbon release is in the form of isoprene, a highly volatile, reactive compound with wide ranging environmental effects. Although most isoprene is attenuated through atmospheric reactions, a significant portion is thought to be degraded by biological processes in soil. In order to investigate the soil isoprene consumption, in situ seasonal measurements of isoprene degradation were performed with atmospherically relevant concentrations. Isoprene spiked microcosms were used to determine the effects of isoprene concentration and the proximity of soils to trees with differential isoprene production on the isoprene-degrading capacity of soil bacterial communities. Bacteria that incorporated carbon from isoprene into their biomass were identified using DNA stable isotope probing (DNA-SIP), enriching bacterial communities on ^{13}C isoprene. Changes in bacterial community composition were analysed in all microcosms, enrichments and DNA-SIP fractions by high throughput sequencing of amplified 16S rRNA genes. The seasonally averaged in situ rate of isoprene degradation was 4.77 nmol m^{-2} h^{-1}, which is lower than previously reported values, with no significant seasonal differences. Also, in microcosms, the isoprene-degradation rate was the same irrespective of the tree species present where the soil sample was taken. There was no significant change in isoprene degradation rate between communities dependant on proximity to natural isoprene sources, and numerous organisms involved in isoprene degradation were identified; with *Rhodococcus* species being the primary degraders at high concentrations of isoprene, but potentially less relevant in the environment than previously assumed, and *Methylobacterium* potentially having a role in isoprene degradation in the terrestrial environment.

2.2 Introduction

Isoprene (C_5H_8) is makes up 1/3 of the global VOC release into the atmosphere. Approximately 600 million tonnes of isoprene are produced each year, with the vast majority coming from plants (Guenther et al., 2006). In leaves, isoprene is proposed to stabilise the thylakoid membrane, allowing plants to better resist thermal shock (Sharkey et al., 2008). It also has a role in protection against oxidative stress (Vickers et al. 2009). However when highly reactive isoprene reaches the atmosphere it has numerous fates, with most of the isoprene attenuation taking place through reactions with atmospheric gasses (Fan & Zhang, 2004); one of which is to combine with nitrogen oxides resulting in the production of tropospheric ozone, which can have significant adverse effects on the health of plants and animals (Ashworth et al., 2013; Jenkin & Clemitshaw, 2002). Isoprene can also react with hydroxyl radicals, resulting in fewer radicals that would otherwise react with methane, extending its residence time, thus enhancing global warming (Collins et al., 2002). Isoprene can also lead to the formation of secondary organic aerosols (SOA) (c. 1.6 Tg SOA yr^{-1}), which can potentially increase radiative forcing by acting as cloud seed nuclei through cloud processing (although this can be considered condition dependant, as inhibition of aerosol nucleation events by isoprene has also been observed), leading to cooling effects,. SOAs can also exasperate respiratory disease in humans (Kiendler-Scharr, et al., 2009; Turpin et al., 2005). Due to the reactivity of isoprene, despite the high production, isoprene concentrations in the atmopshere tend to be in the low parts per billion, going down to the low parts per trillion (e.g. above marine environments) and maximums of 10 ppb (e.g. above tropical rainforests (Baldocchi et al., 1995;

Kesselmeier et al., 2000; Lewis et al., 1997). Most isoprene is attenuated by a suite of atmospheric reactions, however it is theorised, and there is some evidence to suggest that, similarly to most other volatiles, a proportion of atmospheric isoprene ends up degraded biologically in soils (Tyler et al., 1991; Pegoraro et al., 2005, Cleveland & Yavitt, 1997,1998; Trudgill et al., 1994),. A number of isoprene-degrading bacteria have been identified and isolated, including *Nocardia*, *Rhodococcus*, *Alcaligenes*, *Gordonia*, and *Mycobacterium* species, and initial work into *Rhodoccocus* AD45 has elucidated an isoprene degradation pathway (El Khawand et al., 2016; Crombie et al., 2016; Vlieg et al., 1998; van Ginkel et al., 1987; Ewers et al., 1990).

Although some work on biological isoprene degradation in the environment has already been done, it is understudied compared with the degradation of other hydrocabons, such as methane. Research into isoprene degradation has so far mainly centred on a few small scale studies (Acuña Alvarez et al., 2009; Cleveland & Yavitt, 1998; Cleveland & Yavitt, 1997; Pegoraro et al., 2005).

The initial attempts to characterise biological isoprene degradation rates led to a global estimate of 20.4 MT y^{-1}, or around 3 to 4% of isoprene production being degraded by biological processes in soils (Cleveland & Yavitt, 1997), indicating that soils represent a small but potentially important sink of isoprene. This estimate was based mainly on ex situ rates on soils from a number of locations in the USA and its incorporated territories, with isoprene added at 508 ppbv, as well as in situ evidence using 385 ppbv in static field chambers, giving an isoprene degradation rate of 1428 nmol m^{-2} h^{-1} (Cleveland & Yavitt, 1997; 1998). Other work using increased concentrations of isoprene (100-1000 ppb) indirectly hinted at much higher estimates (14500 to 61200 nmol m^{-2} h^{-1});

(Pegoraro et al., 2005). However, these isoprene concentrations greatly exceed measured values of isoprene in the atmosphere, which are generally less than 10 ppb (for example: Alves et al., 2014; Kanawade et al., 2011; Kesselmeier & Staudt, 1999; Sharkey & Yeh, 2001). The estimates so far for global isoprene biological degradation has been based on rates determined at these comparatively high concentrations compared with more realistic atmospheric concentrations (i.e. 3 ppbv) (Cleveland and Yavitt, 1997). Since then, an *in situ* test, using the atmospheric isoprene present at the time of sampling has demonstrated an isoprene-degradation rate of 2.0 nmol m^{-2} h^{-1} (Gray et al., 2014), which is significantly lower than previous estimates. It is notable that the isoprene-degradation rates of soils increase as the concentration of isoprene increase, and are higher in studies where the atmospheric reactions, temperature, or moisture content are controlled, (as they were in the experiments leading to the original estimate of a c. 20 Tg global isoprene sink) (Gray 2015, Pegoraro, 2005, Cleveland and Yavitt, 1998, 1997). This ability of soil microbiota to degrade isoprene at higher rates than they would be exposed to atmospherically may reflect the fact that isoprene is produced in soil as well as from vegetation. Although there is no direct measure of soils producing isoprene (although partially heat sterilised soil produces low concentrations of isoprene, probably due to *Bacillus* survival and growth (data not shown)), isoprene is produced from household waste/compost (Mayerhofer et al., 2006). Also, soil dwelling bacteria (Kuzma et al., 1995) and fungi (Back et al., 2010) produce isoprene. However, those few plants tested, including poplar, don't produce isoprene in their roots (Cinege et al., 2009) (at least naturally, inserting isoprene synthase and inducing plant wide overexpression is an exception

(Loivamaki et al., 2007)). A microbial source for isoprene in the soil could lead to degrading microbes adapted to locally high concentrations.

Due to the difference between experimental and actual conditions, it is likely that the original estimates for isoprene consumption are overestimated; however with only one study indicating a lower estimate, there is still uncertainty as to how important soils are to global isoprene flux.

As with most biological processes, there are expected to be a multitude of factors influencing isoprene degradation rates in soils. So far, work on these have been focussed on those factors required to service models and forecasts of global isoprene degradation rates; pH, forest type, moisture content, depth, and CO_2 (Cleveland & Yavitt, 1998; Cleveland & Yavitt, 1997; Pegoraro et al., 2005). Although this gives a coherent, if low resolution, understanding of the effects on these factors on isoprene-degradation rates, and has been applied to the observable atmospheric levels of isoprene to inform our soil isoprene flux estimates, seasonal effects have not been explicitly studied (Cleveland & Yavitt, 1998; Cleveland & Yavitt, 1997; Pegoraro et al., 2005). The seasonal differences in isoprene production from the main source, broadleaf trees, results from leaf loss and differences in temperature and insolation (Guenther et al., 2006; Monson & Fall, 1989), as well as the isoprene emissions being related to season (low in spring and autumn, compared to summer) and temperature (Harley et al., 1996; Monson et al., 1994), and therefore you would expect soil communities to be better adapted to degrading isoprene in summer months, both because there is potentially more isoprene available, and because their metabolic activity should be higher in warmer soils. However, isoprene is still present in the atmosphere in autumn (Li & Wang, 2012), and

the atmospheric factors also change. For example, there is a noticeable decrease in hydroxyl (OH) radicals in autumn (Atkinson, 1990), which could conceivably lead to longer residence times for isoprene. Therefore, it would be useful to understand seasonal effects on isoprene degradation rates, to better understand isoprene flux, and the potential future impacts of a biological sink. Additionally, it would be useful to understand seasonal effects from the perspective of isoprene degradation; as temperature and soil water content affects isoprene degradation rate, there could be a seasonal effect (Cleveland & Yavitt 1998, Pegoraro et al., 2005). Likewise, with the effect of production source on soil degradative capacity; as the foliage of some trees is expected to be the main source of isoprene to soils, with roots not producing any isoprene, and containing very little isoprene synthase (although soil isoprene production levels are unknown) (Cinege et al., 2009), and as isoprene has a very short atmospheric lifespan (Atkinson & Arey, 2003), we could reasonably expect that soils near to isoprene-producing trees would contain bacterial communities able to degrade isoprene more readily. Should there be significant effect on the ability of soils to degrade isoprene from the production capacity of nearby tree species, then this should be factored into future models of isoprene degradation. If there is not a significant effect of proximity to isoprene-producing trees and no noticeable seasonal effects on isoprene degradation rates of soil microbiota, then this would raise questions about why soil communities, which are highly competitive environments (Hibbing et al., 2010), are primed for isoprene degradation, when the known sources are absent.

The study of the effects of isoprene on soil communities is important in understanding the global isoprene cycle because: (i) soil microbiota, with an

estimated 2.6×10^{29} soil bacteria globally, and up to 10^{10} bacteria (across up to 38,000 species) per gram of soil, are an important part of atmospheric gas transfer (Whitman et al., 1998); (ii) how the communities change can be used to inform models of isoprene flux from a bottom up approach (i.e. calculating global changes in biological isoprene degradation rates based on microbial community changes under the changing conditions); and (iii) currently most of the mechanistic understanding we have of isoprene degradation is based on a few culturable organism, yet most microbes have not been cultivated (Rappé & Giovannoni, 2003). The identification of isoprene degraders is the first step in building upon this knowledge base to develop a clear understanding isoprene degradation.

Our current knowledge of the bacterial species involved in isoprene degradation is limited, with the following genera having been linked to isoprene degradation: *Rhodococcus*, *Nocardia*, *Arthrobacter*, *Gordonia*, *Mycobacterium*, *Leifsonia, Alcaligenes* and *Alcanivorax, Pseudomonas, Alcaligenes, Klebsiella* (van Hylckama Vlieg et al., 1999; Acuña Alvarez 2009; Ewers et al., 1990; Srivastva et al., 2015; van Ginkel et al., (1987). In soils, Gray et al. (2015) also suggested an increase in the relative abundance of the actinobacterial genus *Sporichthya* up to 40-fold with isoprene over 8 days at 200 ppb, compared to incubation with no added isoprene, which looks like it could be significant, as well as small changes at the phyla and Proteobacteria class level, for which significance is claimed (although $^1/_3$ of those have too low correlation coefficients for averages against concentrations for P to be <0.05).

Recent work conducted by our project partners at UEA has also indicated that direct incorporation of isoprene derived carbons into DNA may occur in marine

environments in *Rhodococcus* and *Mycobacterium* (Johnston, 2014). In soil environments, low resolution family level experiments identified a number of families involved, suggesting isoprene degradative ability in a diverse range of groups, and further supporting the *Rhodococcus* genus containing isoprene degraders (El Khawand, 2014). However, in many of these analyses, there is often limited statistical analysis between conditions (Gray et al., 2015; Johnston, 2014; El Khawand, 2014), largely due to the available technology at the time, however limiting confidence in the results.

When the low level of statistical proof is combined with the fact that most of these experiments were performed at significantly higher than environmentally relevant isoprene concentrations, although high concentrations are advantageous in some ways, it is evident that there remains a lot to learn about isoprene-degrading bacteria in the environment.

In this study, a multi layered approach was adopted to investigate soil isoprene degradation; in situ measurements to determine degradation rates, serial enrichments to investigate isoprene degraders with high sensitivity, DNA Stable Isotope Probing (DNA SIP) to demonstrate incorporation of isoprene into DNA, low or realistic concentration experiments to more closely mimic in situ conditions, (in addition to cultivation approaches, which are the subject of Chapter 3).

2.3 Materials and Methods

2.3.1 In situ sampling

In situ degradation rates were investigated in Wivenhoe Park, Essex, U.K. A sample site was chosen in an area with no nearby tree cover. The soil was loamy, undisturbed, and covered year round in a grass moss blend, with few other plants present. The area was flat, and at the top of gently sloped ground, ensuring no excessive water logging. The soil had a year average pH of 4.9, which was representative of the area, when compared with the other sites measured (i.e. not an outlier). As part of the park land, the foliage was subject to regular mowing, and the ground cover had evidence of fowl and rabbit presence. The location details are shown in Table 2.1.

The in situ isoprene degradation rate of soil at each location was measured at four points dispersed throughout 2015 (24/3/15, 18/6/2015, 29/9/2015,10/12/2015).

In situ gas chambers were designed, refined and deployed. Chambers consisted of 2 litre clear polycarbonate bottles (Nalgene validation flasks (Nalgene)), with a diameter of 12.3 cm, and a polycarbonate lid containing a thermoplastic elastomer sealing ring. Through each lid a 25 mm computer fan (Ebmpapst 252 (R.S. components)) was suspended using Dekabon, PTFE coated aluminium tubing (S.P.A. Design), with the wires internalised and powered by an external 12 V battery (R.S. components). Also, through the lid, a sampling port made from a Swagelok 1/4” to 1/8” reducing union (Swagelok) containing a Polytetrafluoroethylene (PTFE) covered silicon septa (Sigma-

Aldrich) was inserted. Joints between the Dekabon tubing and the fan, and lid were sealed with araldite, and allowed to set and stabilise (Figure 2.AX.8).

Three locations, approx. 5 m apart, were selected (and represent the unit of replication). At each location there were two test chambers placed into the soil, one open control (no isoprene addition) and a closed control chamber (with the base not removed). Chambers were prepared with a sharp edge, allowing rotational movement to cut into the soil with minimal disturbance of the central section. A three centimetre insertion was enough to create a sufficiently gas tight seal, and the isoprene addition volumes were adjusted accordingly (due to the small chamber space loss). Chambers were allowed to settle for at least 15 minutes before sealing and addition of isoprene and the inert tracer.

Stabilised isoprene (99%, Sigma) was added from a diluted isoprene stock of known concentration to given an isoprene concentration in the chamber of ~15 ppb, and ~60% 1,1,1,2,3,4,4,5,5,5-Decafluoropentane (DFP) (Sigma) was added to give twice the isoprene volume, as an inert, non-biologically degradable tracer, using gas tight syringes.

Samples were taken immediately after volatile addition, and then at 7.5 and 15 minutes; 50 cm^3 of gaseous chamber content was removed from the chamber using a 100 ml gas tight syringe, and injected into one 500 ml Tedlar bag per sample. The pressure of the chamber was allowed to equilibrate for a few seconds immediately after each sample removal to reduce negative pressure drawing isoprene from the soil. Post sampling, soil was collected from the location of each chamber set, and a bulb thermometer was inserted 3 cm into the ground, and an average temperature taken.

In order to measure the volatile addition below the linear response threshold of the chromatography equipment, it was necessary for volatile concentration equipment and methodologies to be created. The iterative design and testing process is described in Appendix 2.AX.1). In brief, a system was designed where gasses from the Tedlar bags could be drawn through a loop of fine 316 stainless steel tubing using negative pressure. This loop was suspended above liquid nitrogen, with a thermocouple controlled boiler (produced by the University of East Anglia, School of Environmental Sciences) maintaining the temperature of the loop at -170°C. The sample was drawn through at 1 $cm^3\ s^{-1}$, allowing the volatiles to condense on the cold metal. Following this, the tubing was placed under a weak vacuum and sealed. The loop was then heated to >80°C using hot water, and the re-volatilised compounds were withdrawn using a 250 µl gas tight syringe, propelled into the syringe by the negative pressure, and allowing the air movement to proceed from the input side to carry them through. These volatiles were then added directly to a Gas Chromatograph with a Flame Ionisation Detector and the volatile concentrations were measured. This cryotrap normally resulted in over 100× concentration, and the ratio of the inert tracer to the isoprene concentration was used to determine the degradation rates, with each rate being compared with the in group sealed control to remove the effect of in chamber non-biological reactions (without reducing the overall variance), and the starting ratio to remove effects of differing initial concentrations. Rates were determined by the linear regression of time against concentration. Soil pH, was calculated in in a 1:5 dilution of 0.1 mol l^{-1} $CaCl_2$, as described in Cools et al. (2010). Dry weight was calculated by

the difference in weight between a few g of soil before, and after drying at 80°C for at least 24 hours when a stable weight was reached.

Porosity was determined by adding a few grams (weighed before and after) to a 50 ml volumetric flask, filling to the line, and weighing, then heating at 80°C in a water bath for three hours, before allowing to cool, refilling and reweighing, with the difference in weight being equivalent of the space of the soil previously taken up by gas for the amount of soil added (Tan, 2005). Soil characteristics and weather data can for the dates of sampling can be seen in Table 2.SI.1.

2.3.2 Ex situ sampling

2.3.2.1 Sequential enrichments with isoprene

Soil was obtained from underneath the canopies of mature trees, which were arranged in small stands of the same species, in Wivenhoe Park UK. Three sites were chosen that had isoprene-producing species Poplar *(Populus canadensis)*, Willow (*Salix babylonica)* and Oak (*Quercus robur)*, two with non isoprene producers, Birch (*Betula pendula*) and Ash (*Fraxinus excelsior*), and one area of open ground with no tree cover (Table 2.1). For each site, samples were taken between 1 and 2 m from the trunk of a selected tree, and triplicate samples were equidistant from each other around the base of the tree. Leaf litter and other debris were removed. Any short vegetation was removed with care to displace as little topsoil as possible. In the autumn of 2012, approximately 2 x 30 cm^3 of soil was collected aseptically from the top 2 cm, for each triplicate from each site. One sample was used to investigate the physical chemical parameters as in Section 2.3.1, and the other to investigate the

isoprene-degradation potential and bacterial composition. The degradation experiments were set up within 6 hours of sampling.

Table 2.1 Location, Tree type and basic characteristics of isoprene-enrichment soil sample origins.

Site	Species	Location	isoprene producer	Dry mass (%)	pH
Willow	*Salix babylonica*	51.87772 °N, 0.948709 °E	✓	68.45	5.43
Poplar	*Populus canadensis*	51.87799 °N, 0.949446 °E	✓	70.14	6.23
Oak	*Quercus robur*	51.878151 °N, 0.948988 °E	✓	58.76	4.41
Ash	*Fraxinus excelsior*	51.878569 °N, 0.948132 °E	✗	72.85	6.09
Birch	*Betula pendula*	51.87828 °N, 0.950293 °E	✗	50.75	3.78
No Trees	N/A	51.878659 °N, 0.947761 °E	✗	51.87	4.90

pH, Dry mass of soil as mean %, n=3 for Dry mass and pH measurements

Each soil sample was homogenised using a spatula, and 1 g of soil for each replicate for each site was added to three 125 cm^3 glass serum bottles. The serum bottles had been soaked in 5% Decon 90 in distilled water overnight, rinsed thrice with distilled water and baked in an oven at 110°C for three hours. The baked glassware was then rinsed thrice with analytical grade acetone that was allowed to volatize; following which the openings were covered in foil and autoclaved. Autoclaving was at 121°C, 15 psi for at least 15 minutes.

To each serum bottle 9 ml of minimal medium was added, with an additional set of six no soil controls with just 10 ml of medium. The minimal medium comprised of: 0.5 g NaCl, 0.5 g $MgSO_4 \cdot 7H_2O$, 0.1 g $CaCl_2 \cdot 2H_2O$, 1 g NH_4NO_3, 1.1 g Na_2HPO_4, 0.25 g KH_2PO_4, 50 mg Cycloheximide, 10 mg $FeSO_4.7H_2O$, 0.64 mg $Na_2EDTA.3H_2O$, 0.1 mg $ZnCl_2$, 0.015 mg H_3BO_3, 0.175 mg $CoCl_2 \cdot 6H_2O$, 0.15 mg $Na_2MoO_4 \cdot 2H_2O$, 0.02 mg $MnCl_2 \cdot 4H_2O$, 0.01 mg

$NiCl_2 \cdot 6H_2O$, 0.05 mg *p*-Aminobenzoic acid, 0.02 mg Folic acid, 0.02 mg Biotin, 0.05 mg Nicotinic acid, 0.05 mg Calcium pantothenate, 0.05 mg Riboflavin, 0.05 mg Thiamine HCl, 0.1 mg Pyridoxine HCl, 0.001 mg Cyanocobalamin, and 0.05 mg Thioctic acid (phosphates and salts were autoclaved separately, and vitamins were filter sterilised) per litre Milli-Q filtered water (a modification of Fahy et al., (2006)).

Isoprene was added at two different concentrations for each site; to one batch 1 cm^3 of isoprene stock headspace was added, yielding 7.2 $\times 10^6$ ppbv final concentration; to the other batch 0.1 cm^3 isoprene stock headspace was added, yielding 7.2 $\times 10^5$ ppbv final concentration. Additionally, in order to distinguish the effects of isoprene on bacterial community composition from bottle effects, each site/replicate had no isoprene addition (no isoprene controls). Three no soil controls had 0.1 and 1 cm^3 isoprene addition in order to ensure that isoprene degradation was biological. Incubations were performed at 20°C in the dark.

The isoprene stock was made in a similar manner to that of Acuña Alverez et al. (2009), with10 ml isoprene (Sigma >99%, 464953) in a 125 cm^3 washed and sterilised serum bottle, crimp sealed with a Si/PTFE septum, and warmed to 30°C for at least 15 min before use. Isoprene aliquots were taken from the headspace.

Isoprene was measured daily using a 100 µl headspace sample taken through the septa with a glass gas tight syringe with a steel side port needle and injected into the sample port of a Unicam 610 series GC-FID equipped with a 10% Apiezon L CW column, with injector and detector temperatures of 160°C

and a column temperature of 100°C. The GC-FID was calibrated daily by fresh dilutions of isoprene stock, as well as a 100 ppm isoprene standard (Scientific and Technical Gases Ltd.), and had a linear response / concentration relationship through dilutions from 0.1 ppm to 100 ppm (which represents the usable range), and detection, but with a non linear and error prone response under ~70 ppb, and no detection at all under ~5 ppb.

After isoprene concentrations were reduced to less than 20% of no soil control concentrations, 1 ml of the of the soil suspension was added to fresh media, with isoprene added at the same concentration as used in the original microcosm, and 0.5 ml samples were frozen for genetic analysis. This sequential addition was performed two more times, for a total of three enrichments per sample.

Isoprene degradation was reported as the time taken to achieve 80% degradation, with estimation by regression and interpolation of nearby points displaying a trend, where the threshold was passed between testing points. Significance testing was performed using a nested ANOVA in R (R Development Core Team, 2008) (nested tree species within isoprene production status).

For bacterial 16S rRNA gene sequence analysis, DNA was extracted from the soil enrichments using the Griffiths' method (Griffiths et al., 2000), which in brief involves bead beating the sample in Cetyl trimethylammonium bromide (CTAB), phenol chloroform isoamyl alcohol, followed by PEG precipitation), and amplified by PCR using a subset of the methods described by Illumina (Illumina Inc., 2013). Specifically, amplification was carried out on the V3 and V4 region

of the 16S rRNA gene, with the default primer with adapter sequences described by Illumina (Illumina Inc., 2013) (bacterial primer pair S-D-Bact-0341-b-S-17 and S-D-Bact-0785-a-A-21 from Klindworth et al., 2013 with adapter sequences). The first round of PCR was carried out as described by Illumina inc., 2013, except that RedTaq (Sigma Aldrich) was used instead of Kapa Hifi because it increased specificity. In addition, plate spinning in a custom 96 well plate adapted OXO spinner (1351580UK) was performed prior to the PCR. A Bioanalyzer chip was not used at this point, however PCR clean up was performed as described by Illumina inc., 2013 (without MIDI plate shaking), and the second round of PCR was performed as described by Illumina inc., 2013, using the horizontal Illumina indices: N701, N702, N703, N704, N705, N706, N707, N708, N709, N710, N711, and N712, and vertical indices: N517, N502, N503, N504, N505, N506, N507, N508. The second round PCR cleanup was performed as described by Illumina inc., 2013, without shaking. Following this, 5 µl of DNA from each reaction was run on a 2% agarose gel, alongside a pool of the first round products, to ensure the expected sizes and size shifts were present.

Library quantification was performed by diluting the library 1:10, adding an equal volume of Quant-iT PicoGreen dsDNA Assay kit (Invitrogen) preparation, and using a 384 well optical plate, with samples in duplicate, six blank samples, and a standard curve of known concentrations of DNA. The plate was read by a (Biotek Synergy) plate reader set to 485 nm excitation and 520 nm emission fluorescence intensity, and pooling volumes were calculated to maximise DNA (the minimum concentration created with the total volume). The pooled library was analysed in triplicate using a Bioanalyzer 1000 (Agilent) chip and a

LabChip caliper Bioanalyzer (G2938C), following the default DNA 1000 Assay protocols, alongside the initial pool to ensure sufficient quality. DNA was concentrated by collecting the DNA using a PCR purification kit (Sigma; GenElute), running multiple aliquots of DNA binding buffer solution through the same column prior to elution, and DNA quantification was performed using an optical nanodrop, and diluted to 20 ng^{-1} µl.

Sequencing was performed with MiSeq 300 bp paired end reads by the NBAF facility at the Liverpool Centre for Genomic Research. Due to on going issues causing decreased MiSeq read quality, raw reads were obtained for analysis.

In order to analyse these data, a custom pipeline was created. Samples were downloaded pre-separated into sample folders according to the sample list/primer sequences given. Unix based analysis was performed using BioLinux 8 (Field et al., 2006). As suggested in Schirmer et al. (2015), reads were subjected to Bayes Hammer error correction in SPAdes, which can remove 93% of Miseq read errors by hamming and quality adjusted Bayesian subclustering (Bankevich et al., 2012). This approach was used here as a pre-processing step due to an increased error rate (at the time Illumina had a reagent issue, producing poorer quality reads). To ensure that this approach did not adversely affect the analyses, three samples were also ran using the same pipeline but with the original Schirmer et al. (2015) approach, and shown little observable change (a comparison is available in Figure 2.SI.158). Following this, reads were trimmed using Sickle (Joshi et al., 2011), with a quality value limit of 18, to increase data retention from the largely error corrected (from the previous step) reads which were initially carrying lower than desirable quality scores from the initial reads. Following this, paired-end reads were paired using

PANDAseq (Masella et al., 2012), which retains the best bases from the overlapping reads. Sequences were then stripped of extraneous data in the name, and renamed after the sample number, read number (for identification), and with a count identifier set to 1 using GNU sed and awk, pre-emptively solving compatibility issues. These reads were then de-replicated using FastX collapser from the FASTX-Toolkit (Hannon Lab), and concatenated (GNU cat) to the other reads into one file, so that the OTUs would be directly comparable between sites. Clustering was performed de novo using Swarm (Mahé et al., 2014) on the University of Essex Unix-based Genomic Cluster, with d=1, allowing for natural OTU sizing within a fine threshold of stress, and stress central centroid picking, mimicking more closely actual OTU and similarity dispersion than arbitrary fixed threshold methods, which mostly use the first, most, or random methods of centroid picking, and which are not responsive to the underlying phylogenetic community structures when clustering. Chimera checking was performed within vsearch (Rognes et al., 2015) against the Gold database (Pagani et al., 2012). In every case, putative chimeras were below 0.5% of the sequences analysed. Clustered sequence files were converted for DOS compatibility using sed. Classification was performed using the Ribosomal Database Project sequence classifier (Wang et al., 2001), and fix rank level grouping and addition was performed using excel. Bespoke VBA programs within excel were generated, validated and used to: (1) parse the clustered sequence files and generate OTU tables from the Swarm output, (2) draw the applicable centroid sequences from the non chimera list, rarefy against real random number lists to group minima (Haar, 2008), (3) convert to relative abundance, (4) perform unadjusted ANOVAs (using the standard equation

((n∑(Xj- x̄..))/(K-1))/ ((∑((Xij - x̄j)2))/(N-K))), where: x̄.. is the grand mean, x̄ is the group mean, K is the number of groups, N is the total sample size, I is the indexed variable, j is the group variable, and significance attributed when the F statistic passes the threshold for the sample sizes), (5) generate QQ statistics (using the inverse of the cumulative standardised normal distribution, linear regression and calculating the correlation coefficient), (6) perform unadjusted Kruskal-Wallis tests (Using the standard equation; $H=(12/(n(n+1)))\times \sum[^{k}_{i=1}](R_i^2/n_i)-s(n+1)$, where R is the rank sum for sample i to k and the p value estimated using the excel chi-distribution function), as appropriate, and to (7) generate graphs, for every OTU in every test and condition, as well as with the RDP fix rank grouped data, in addition to performing automated data subsetting and R code generation for between source permanovas performed in using Adonis within the Vegan package (Oksanen et al., 2015) in R (R Development Core Team, 2008), and analysed in excel.

Multivariate tests were preformed using the manyglm function in the mvabund package in R to fit negative binomial models with the *P*-values calculated using 1000 resampling iterations via PIT trap resampling, and testing of models by anova. Plots of Dunn and Smyth residuals and plots of the mean variance of the model were used to confirm model assumptions (Wang et al. 2012).

Species level identification was performed using Blast (Altschul et al., 1990), for similar strain name identification, followed by looking up type strains on LPSN (Parte, 2014), and closest type strain identification by phylogenetic tree generation using MEGA 6, MUSCLE, and the Jukes and Cantor and Neighbour Joining methods with 1000 times bootstrap statistics, for analysis at the OTU level (Jukes and Cantor, 1969; Saitou and Nei, 1987; Felsenstein, 1985). At the

site level, analysis was performed using the Non Metric Multidimensional Scaling (NMDS) wrapper in R vegan (metaMDS), with custom graph generation, point identification and analysis in Excel, using VBA.

2.3.2.2 DNA Stable Isotope Probing of soil with isoprene.

Soil was taken from different points in the same Willow and Birch locations as used in the previously described enrichments. Soil was sampled in Summer 2015 (2 June 2015), and soil microcosms were set up in the same manner as described in 2.3.2.1, but the experimental setup and isoprene addition differed. Willow and Birch soils were incubated in minimal media, with either 1.0 ml from stock headspace of ^{12}C isoprene or ^{13}C isoprene (Fully labelled ^{13}C isoprene kindly provided by DuPont Industrial Biosciences: California, USA), resulting in an initital headspace isoprene concentration of 7.2×10^6 ppb (0.1 M), or no isoprene. The ^{12}C stock was made as in section 2.3.2.1, whereas the ^{13}C stock was made by just exceeding the minimum quantity of isoprene required to saturate the headspace (~175 µl) of the 125 ml bottle, owing to the limited quantity of isoprene available. All samples were confirmed to have the required concentration at the time zero GC-FID measurement. Isoprene degradation was monitored by GC-FID, and samples were destructively sampled for DNA extraction (as described in Section 2.3.2.1) after one and four days. Caesium Chloride (CsCl) ultracentrifugation (Sorvall Discovery 90SE with vertical rotor, 44100 rpm, under vacuum, with acceleration and deceleration program 1) was performed on DNA from the ^{13}C and ^{12}C isoprene incubations from the day 4 samples from the Willow soil enrichment.

DNA was added to 6 ml CsCl tubes (Sorvall 79273 6.0 ml polyallomer ultracrimp), for DNA-SIP according to Neufeld et al. (2007), without ethidium

bromide and increased spin times of 48 hours. The DNA containing CsCl was fractioned into 12 × 0.5 ml fractions (to the nearest drop), in the same manner as described by Neufeld et al. (2007) except that a peristaltic pump with sterile water was used to control the gradient exit. Gradient formation was confirmed using an optical refractometer to measure the refractive index, which was checked against a standard curve of CsCl, created by measuring densities by weighing defined volumes and measuring with an optical refractometer.

End point PCR amplification using the Illumina first round primers and conditions, with agarose gel separation (1.5%, 1 hour, 110 V, Fermentas SM 0333 ladder), followed by ethidium bromide staining and UV visualisation, were performed to ensure an increased spread of DNA towards the heavy fractions was present in the ^{13}C fractions compared to the ^{12}C fractions.

Next generation sequencing was performed in the same way as in Section 2.3.2.1, on the fractions 1 (heaviest) to 10 (near-lightest) of the ^{12}C and ^{13}C isoprene incubations, as well as the pre-incubation soil community.

Bioinformatics analysis was performed in the same way as in Section 2.3.2.1, in a separate batch, alongside the low concentration experiments, however with an additional detrended correspondence analysis performed due to the gradient nature of the density gradient allowing for improved point separation by this method. To see the difference in relative abundance caused directly by isoprene degradation, and leading to DNA incorporation, statistics were applied to, and graphics were produced from ^{13}C relative abundance, net of the ^{12}C relative abundance for each fraction (i.e. the difference between the same fraction in each carbon weight set).

2.3.2.3 Low concentration experiments

The same Willow and Birch soil samples used in Section 2.3.2.2 were used in the low concentration experiments. Soil (0.5 g) was added to 125 cm^3 serum bottles without any media addition. Isoprene was added daily (except for weekends) at 0, 15 and 150 ppb concentrations for 23 days, resulting in 16 separate applications. GC-FID was used to confirm degradation of the isoprene in the samples, however, under 100 ppb, the response of the GC-FID has high internal error, meaning that isoprene concentrations at the 15 ppb level are unlikely to be accurate. Moisture levels in the soil were kept constant by saturating the headspace after every isoprene replenishment. Approximate calculation of saturated vapour density using the <40°C approximation curve: $VD = 5.018 + 0.32321\ T_C + 8.1847\times10^{-3}\ T_C^2 + 3.1243\times10^{-4}\ T_C^3$ (Relhum), calculated using the estimated temperature and humidity from the BBC weather, gave the volume required for addition of sterile water to the side of the vessel to bring the internal air to the point of water saturation. Incubation was at 20°C in the dark. After the incubation period, DNA was extracted, sequenced and analysed as in Section 2.3.2.1.

2.3.2.4 Amalgamated analysis

In order to make sense of the data produced by the different experiments, a funnel was applied to the results.

First, in order to reduce the chances that significance isn't being noted due to random chance, amongst a high number of statistical tests; inclusion criteria were applied requiring multiple significant enrichments at a relative abundance where starting points of that (or other) OTU fix rank groups exist (i.e.; a species cut off dependant on the lowest relative abundance of an identified genus level

OTU within the experimental data set). Groups were also included for analysis where isoprene-degrading isolates were identified in Chapter 3 of this thesis. These OTU groups were investigated for OTUs which showed enrichment within them. These isoprene related OTUs and groups were then individually scrutinised for evidence of isoprene degradation ability, direct usage of isoprene through SIP, and if any environmentally relevant effects through media free low concentration experiments, and related to current knowledge. The representative sequences from each OTU level group were added to a phylogenetic tree, as in 2.3.2.1, with grouping of very similar, independently generated OTUs evidencing the robustness of the analysis approach.

2.4 Results:

2.4.1 The rate of isoprene degradation by microbes within soils is at the lower end of previous estimates, with negligible seasonal effects

Gas chambers were placed on soil from a treeless area of Wivenhoe Park, UK, and spiked with isoprene and decaflouropentane on four days, each in a different season in 2015. Volatile concentrations were measured, and used to estimate a rate of isoprene degradation. The average isoprene-degradation rate was 4.77 nmol m^{-1} h^{-1}, and although there was a small effect of season on isoprene degradation rates, with rates lowest in December and highest in June, this was not significant ($F_{3,17}$ = 0.715 P = 0.56, Anova) (Figure 2.1).

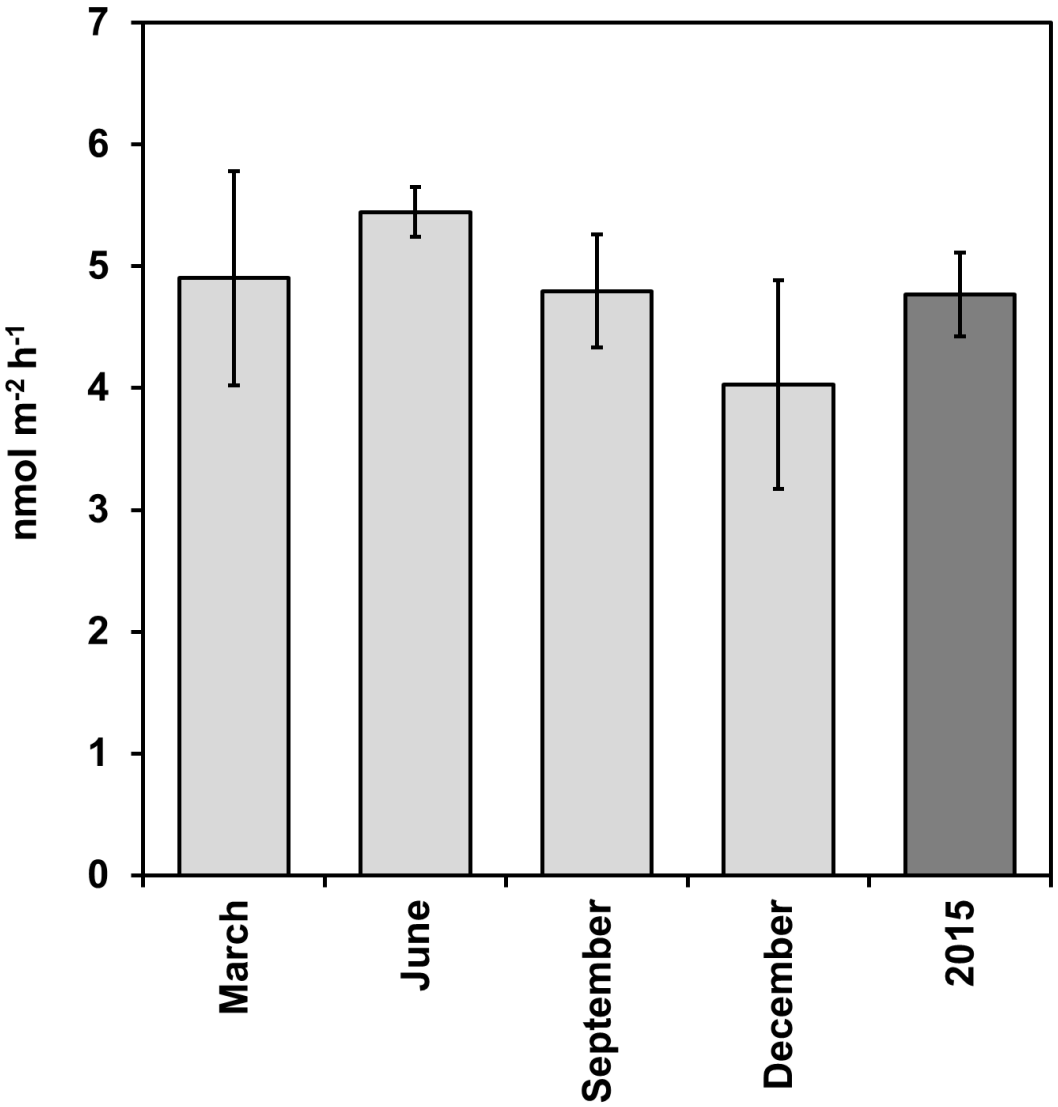


Figure 2.1 Isoprene degradation rates for soil at Wivenhoe Park, UK, (51.878659 °N, 0.947761 °E) at four points throughout 2015. Three groups of four transparent polycarbonate chambers with air circulating fans suspended on Teflon coated aluminium, with Si/PTFE septa mounted in stainless steel, were deployed in the field approximately 5 m apart, between 9 and 12 am. In each group, two test chambers and a sealed control was spiked with 15 ppb isoprene and Decafluoropentane, and one was left without. Chamber air was sampled immediately and every 7.5 minutes, cryo concentrated and analysed by GC-FID. Ratios of isoprene to decafluouropentate were expressed as a fraction of the starting ratio, net of in group control, and extrapolated to m^2. The rate was determined by using the average rate, or linear regression where final time point had no isoprene. n=6,4,6,5 (for 24/3/15, 18/6/2015, 29/9/2015, 10/12/2015 respectively) "2015" is year average for successful tests. Error bars are ±SE.

2.4.2 Microbes within soil degrade isoprene rapidly, but rates are not significantly affected by proximity to isoprene producing trees.

Soil from under isoprene-producing Poplar, Willow and Oak trees, from under non-isoprene-producing Birch and Ash trees, and from a treeless area, were incubated with isoprene to determine the isoprene-degradation rate, and the effect of location on it.

Isoprene was degraded rapidly in the first enrichments, with 80% degradation within 80 h (average of about 55 h) at the lower isoprene concentration (7.2 $\times 10^5$ ppb; Figure 2.2), and 160 h for the higher isoprene concentration (7.2 $\times 10^6$ ppb; Figure 2.3) (note that in Figure 2.2 and Figure 2.3, due to the sampling regime, times taken for degradation are to within one day (depending on exact sampling time), therefore where no error bars were present this was due to degradation of all replicates between testing).

However, there was no significant difference between isoprene producing and non-producing groups ($F_{6,63}$ =0.57 P= 0.75), nor between the different samples locations ($F_{24,63}$ =1.33 P= 0.18), nor between enrichment levels ($F_{2,63}$ =0.48 P= 0.65), and there was significant difference between concentrations ($F_{3,63}$ =36.7 P= >0.001) for the time taken for 80% degradation (nested ANOVA).

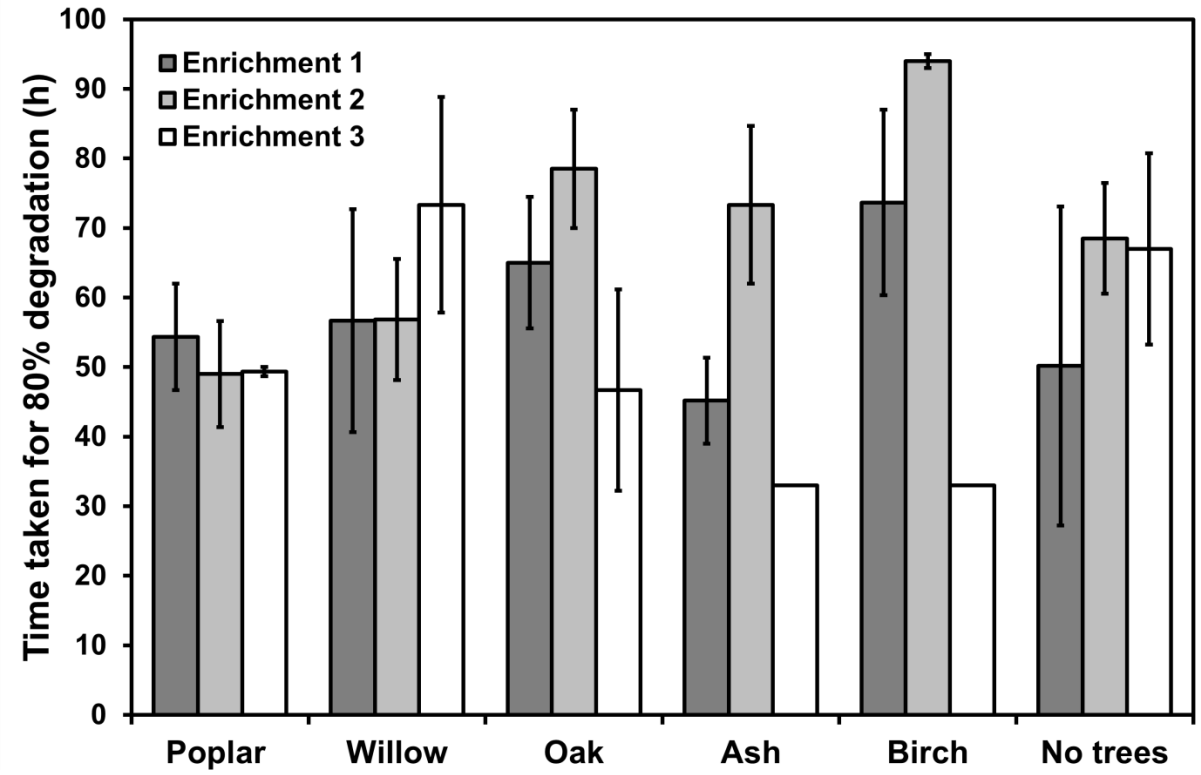


Figure 2.2, Time for 80% isoprene degradation for microcosms containing 1 g of soil from under Poplar, Willow, Oak, Ash, and Birch trees, and an area with no trees, incubated with 0.1 ml of saturated isoprene headspace, at 30°C (final conc. 7.2 ×10^5 ppb), and daily sampling, n = 3, error bars are SE.

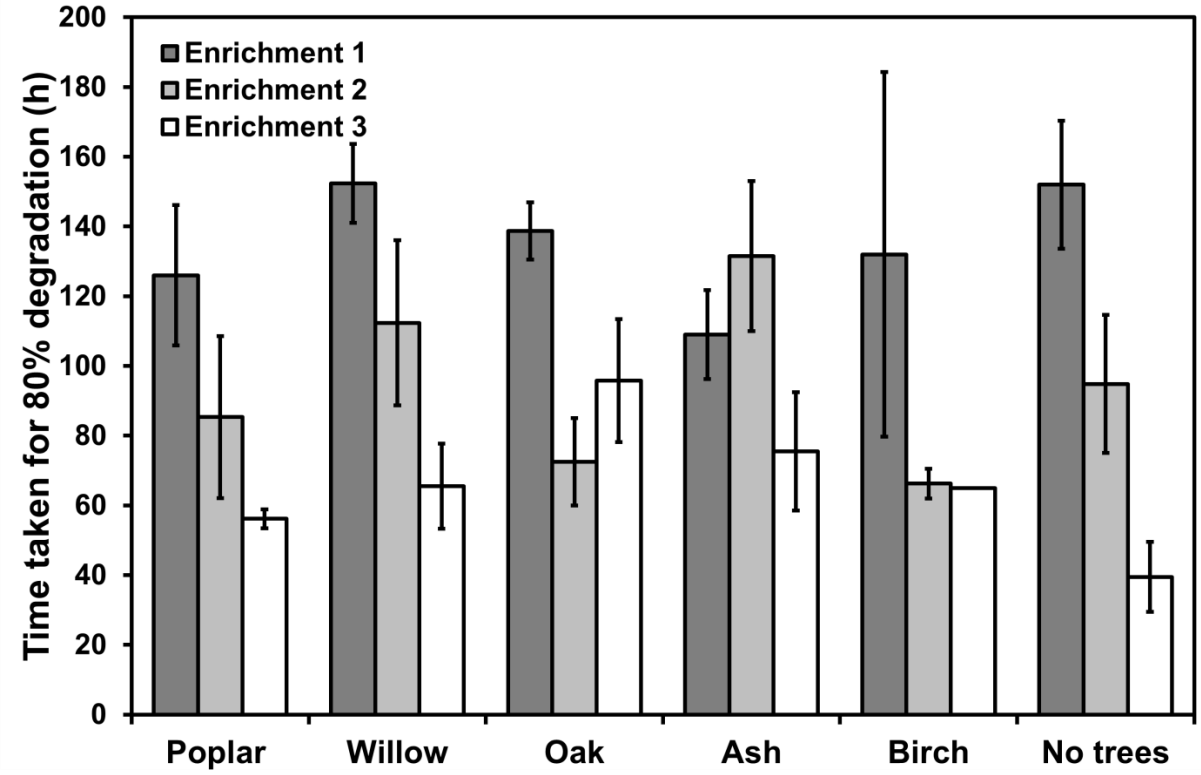


Figure 2.3, Time for 80% isoprene degradation for microcosms containing 1 g of soil from under the canopy of Poplar, Willow, Oak, Ash, and Birch trees, and an area with no trees, incubated with 1.0 ml of saturated isoprene headspace, at 30°C (final con. 7.2 ×10^6 ppb), and daily sampling, n = 3 (n=2 for Birch enrichment 3), error bars are SE.

In general the rates were more higher in the later enrichments, especially for the higher concentrations of isoprene ($p<0.05$ for 0.1 (7.2 ×10^5 ppb isoprene) with Birch, and 1.0 (7.2 ×10^6 ppb isoprene) with all samples (comparing first to third enrichment) by ANOVA), demonstrating increasing isoprene-degrading capability of the enrichments. Graphs portraying raw data, showing change in isoprene concentration in individual triplicates as a percentage of the no soil control are in supplementary information (Figure 2.SI.2).

2.4.3 Sequential isoprene enrichment causes significant changes in bacterial community composition

Metagenetic amplicon sequencing was performed on DNA extracted from soil enrichments sequentially enriched by isoprene to investigate the effect of isoprene on bacterial community composition. Non metric multidimensional scaling (NMDS) was applied to the OTUs, with 2D plotting to visualise community changes.

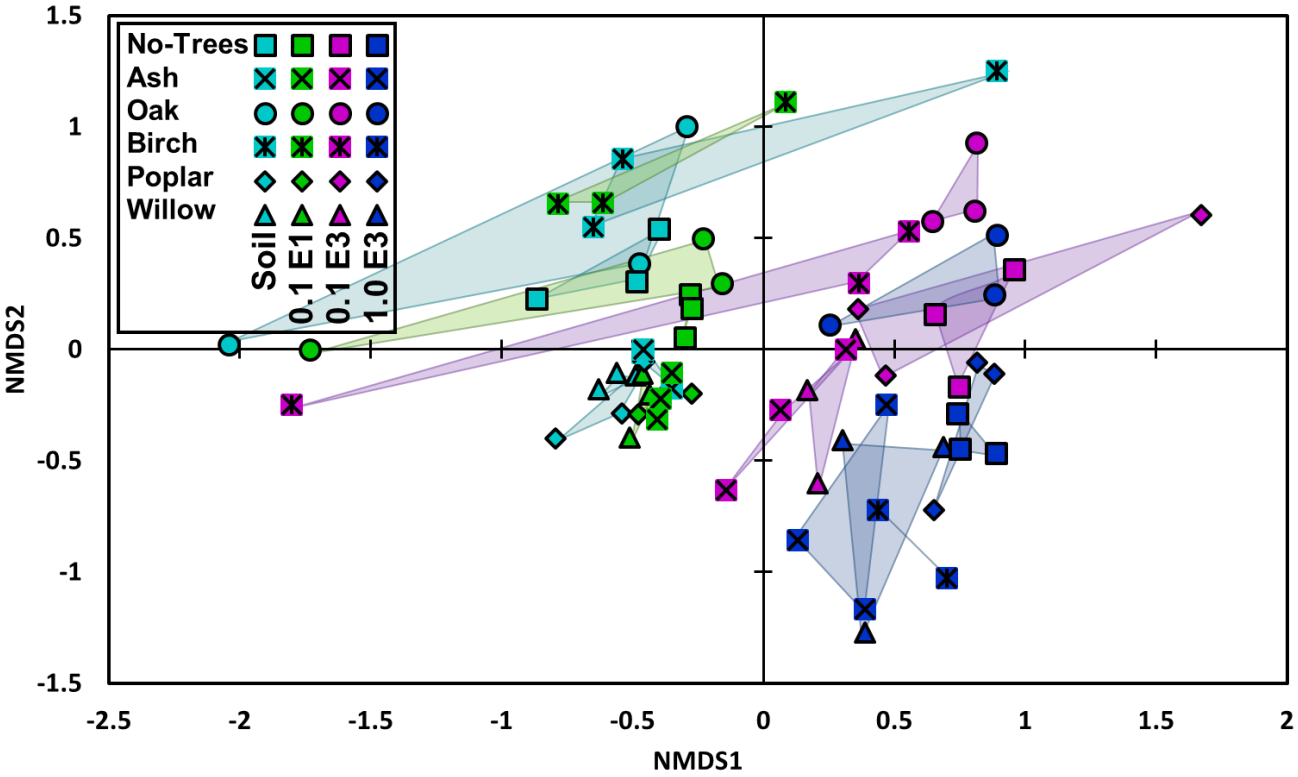


Figure 2.3 Soil enrichment NMDS stress 0.1478054. Soil = Pre-enrichment communities, E1 = First enrichment and E3 = third enrichment. 0.1 = addition of 0.1 ml saturated isoprene headspace (7.2 ×10^5 ppb), 1.0 = addition of 1.0 ml saturated isoprene headspace. n=3 (n=2 for Birch E3 1.0). Figure 2.SI.3 contains the distribution of significances between the enrichments. Significant effects of tree type ($F_{5,65}$ = 1.82 P= 0.021), concentration ($F_{2,68}$ = 18.2 P= >0.001) and enrichment level ($F_{2,68}$ = 21.7 P= >0.001) were observed (permanova).

The original samples varied in the degree of heterogeneity between triplicates. For example, bacterial communities from soil beneath ash, poplar, willow and no trees were comparatively homogeneous, whereas oak and birch were more heterogeneous. The 0.1 E1 enrichments showed a similar pattern, but there is a general shift in bacterial community composition between the original communities and the 0.1 E1 enrichments (significant for ash, willow, and no trees; Figure 2.SI.3) and the degree of heterogeneity is less than in the original samples. Isoprene addition led to clear formation of clusters based on enrichment level (Figure 2.3). Notable is the higher the level of enrichment, the

more the communities move to the bottom right of the NMDS plot, showing that enrichment level affects community composition. Seemingly the communities from different original sources do not cluster as densely as communities from different enrichments, suggesting that isoprene, as a factor, is more important in the make up of the community structure.

The pre-enriched soil and the 7.2 $\times 10^5$ ppb enrichments (Soil, E1) seem to group, and the late (E3, H3) enrichments also form a group, suggesting that the number of sequential enrichments is more important in determining the final community than the level of isoprene the communities are exposed to.

2.4.4 Isoprene use causes demonstrable shifts in bacterial community structure

The direct effect of isoprene on bacterial community structure was investigated through DNA SIP, effectively repeating the willow 1.0 E1 enrichment in the previous study. Carbon from heavy, ^{13}C-labelled isoprene incorporated into bacterial DNA allowed the separation of DNA of isoprene users from other bacterial DNA, with identification through amplicon sequencing.

The willow soil enrichments degraded more than 80% of the isoprene within 4 days (Supplementary information Figure 2.SI.4). Stable isotope probing was successful with clear linear relationships between fraction and density for both ^{12}C and ^{13}C isoprene (Supplementary information: Figure 2.SI.5), and increased PCR product in the high density fractions enriched with ^{13}C compared to ^{12}C isoprene (Supplementary information: Figure 2.SI.6). The initial community was diverse, with around 50% of OTUs under 0.5% relative abundance, with none

above 10% (the original birch soil community is shown for comparison; Supplementary information Figure 2.SI.7).

Detrended correspondence analyses of DNA SIP fraction communities showed: (i) starting communities were very similar, clustering tightly with many mid-low density fractions, (ii) the communities showing the greatest dissimilarity were a mixture of ^{12}C and ^{13}C samples (Figure 2.4), and (iii) the community composition causes increased dissimilarity between the ^{13}C and the ^{12}C labelled DNA as density increases (Figure 2.6), demonstrating that the addition of ^{13}C isoprene lead to incorporation of the heavy carbon into the DNA of a subset of the community which degraded the isoprene. Although the detrended correspondence analysis plot has increased central DCA2 axis variability, this was shown to not be due to third axis distortion (a common issue with DCA, causing a diamond shape), because NMDS of the same dataset created the same overall shape, and similar (but less precise) clustering (and a clear directional heavy/light split outside the initial groupings; Figure 2.SI.8). Graphics depicting the significant differences between and among the different ^{12}C and ^{13}C fractions, pre-enriched soil, and pre-SIP enrichments as determined through permutational manovas are in supplementary information Figure 2.SI.9, with notable significant differences between equivalent density ^{13}C and ^{12}C fractions.

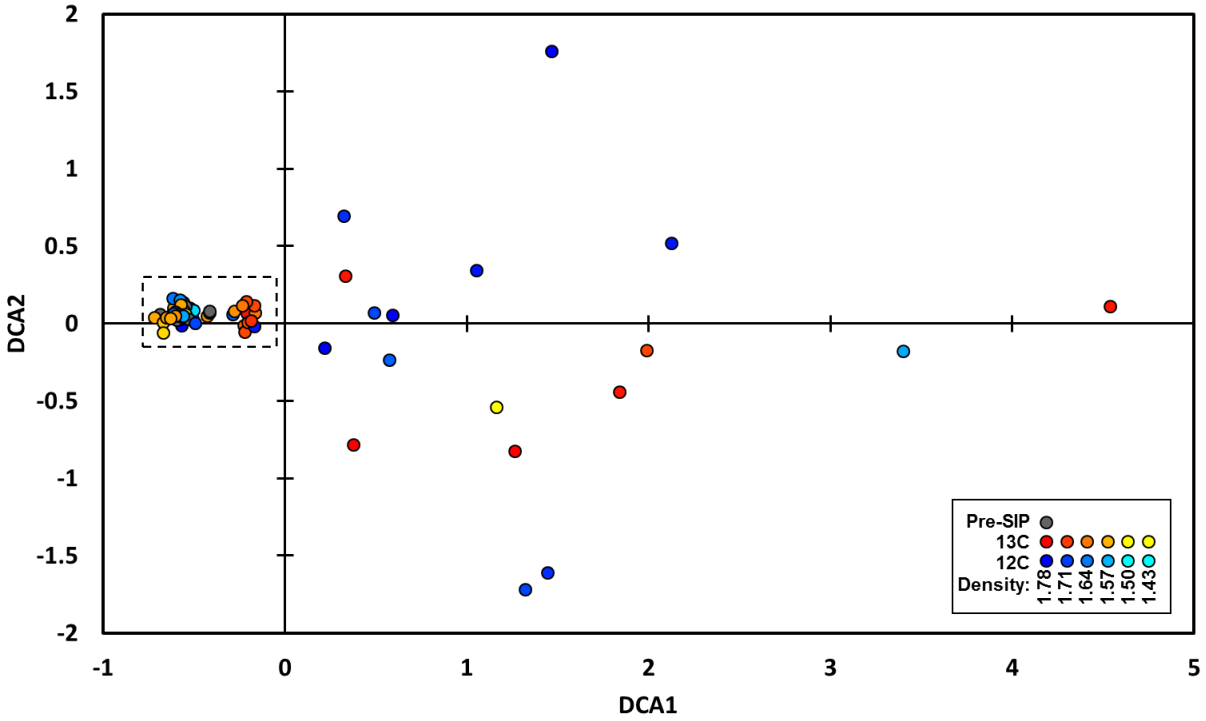


Figure 2.4 Detrended Correspondence Analysis of DNA SIP of soil data showing changes in bacterial community structure for different densities of DNA after enrichment with ^{12}C and ^{13}C isoprene. Points represent fractions from ultracentrifugation. Fill colours yellow to red represent low to high CsCl fraction density with ^{13}C, colours turquoise to blue represent low to high CsCl fraction density with ^{12}C, and Grey represents the starting communities. The data in the box are shown in Figure 2.5. Significant changes in community structure were observed through the CsCl density gradient (a proxy for incorporation of ^{13}C from the ^{13}C isorpene into the DNA) ($F_{10,58}$=2.32 P = >0.001) and with carbon type ($F_{1,67}$=2.48 P = 0.03).

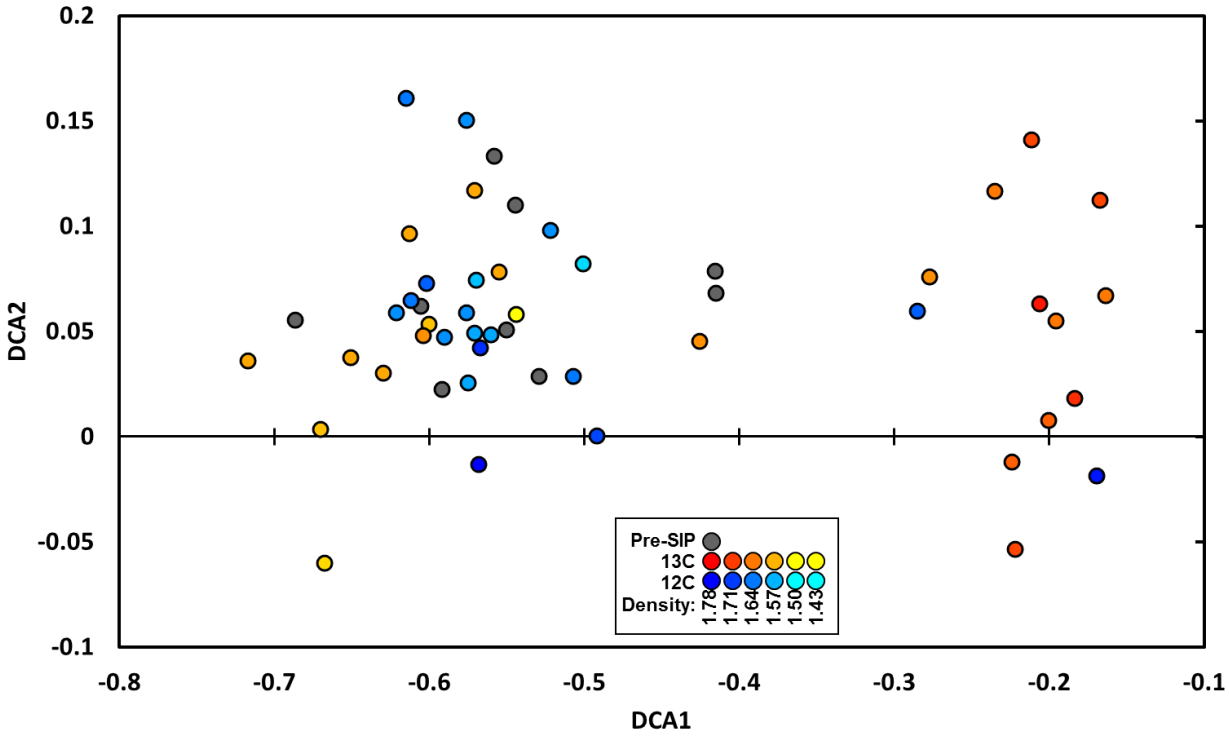


Figure 2.5; Zoomed in section of Figure 2.4 - a Detrended Correspondence Analysis of DNA SIP of isoprene-enriched soil. Points represent sites. Fill colours yellow to red represent low to high CsCl fraction density with ^{13}C, colours turquoise to blue represent low to high CsCl fraction density with ^{12}C, and Grey represents the communities before density gradient centrifugation.

2.4.5 At realistic concentrations of isoprene exposure, enrichment leads to visible but insignificant shifts of bacterial community structure.

Soil from under birch and willow trees was subjected to multiple rounds of isoprene addition at concentrations that are more representative of those found in the environment. NMDS was used to visualise differences in bacterial community structure.

The 150 ppb isoprene concentrations decreased to less than 10% of the starting concentration within 24 hours consistently, and occasionally isoprene was visible in all microcosms (including no isoprene-addition), but under the

linear detection limit (an example, showing the degradation of isoprene with different starting concentrations is shown in Supplementary information Figure 2.SI.10).

Pre-enrichment soils and 0 ppb enrichment soils clustered closely, indicating that any differences in bacterial community structure due to differences in experimental conditions and the original environment were irrelevant compared to those caused by soil location or enrichment level (Figure 2.7). Presence of isoprene seems to cause negative movement on the NMDS2 axis, although insignificantly, and with no signs of convergence of enrichments, and seemingly, a larger effect on the Birch than on the Willow samples. There is a notable increase in heterogeneity in the birch soil enrichment with 150 ppb isoprene. The most obvious differences were along the NMDS1 axis, separating Willow from Birch. However few groups shown significant enrichment with low concentrations of isoprene.

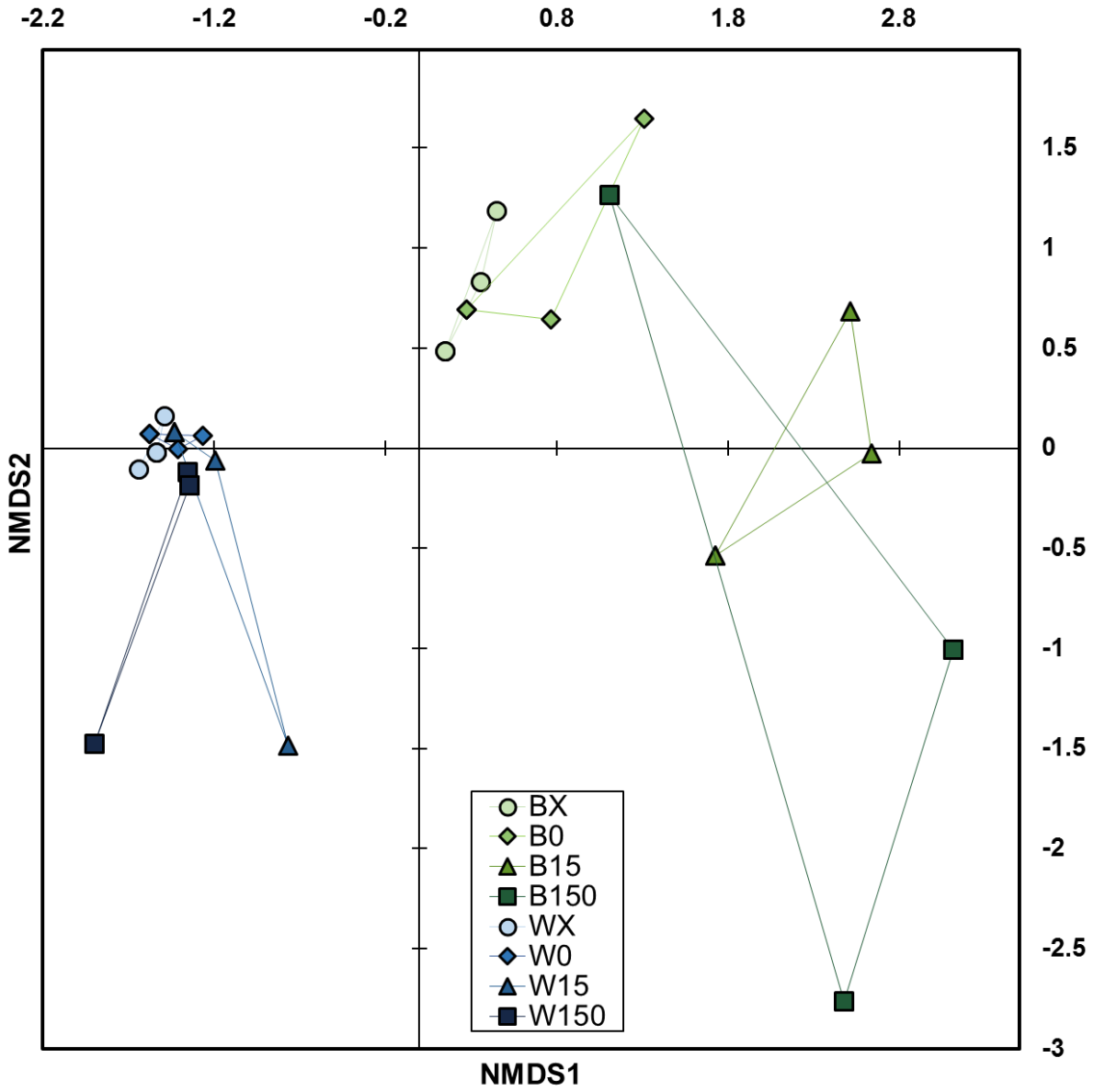


Figure 2.6 Non metric multidimensional scaling plot of bacterial communities of willow and birch soil incubated with 0, 15 and 150 ppb isoprene. B = Birch, W=Willow, X= before incubation, 0 = incubated with 0 ppb isoprene, 15 = incubated with 15 ppb isoprene, 150 = incubated with 150 ppb isoprene. Samples had humidity and isoprene readjustment 16 times in a 23 day incubation prior to sampling. Stress = 0.08987834. No significant effects of isoprene concentration ($F_{3,20}$ = 0.98 P= 0.47), significant effects of tree type ($F_{1,22}$ = 13.8 P= >0.001) (permanova), n=3.

2.4.6 Bacterial OTUs potentially involved in isoprene degradation in soils

To investigate the bacteria enriched in the presence of isoprene, and thus potentially involved in isoprene degradation or probably involved in isoprene

degradation (in the case of SIP experiments), bacterial OTUs were identified with: significant enrichment with isoprene, increased abundances of heavy DNA in isoprene DNA SIP enrichments, and indications of enrichment in low isoprene concentration experiments.

OTUs with centroid sequences classified as *Rhodococcus* were significantly enriched in 11 of 18 enrichments, across the six soil origins (Figure 2.7). At the highest level of enrichment (E3), with the highest isoprene concentration (7.2 $\times 10^6$) *Rhodococcus* often made up over 80% of the bacterial community, reaching over 95% in the third high concentration enrichment of willow soil. Differences between the ^{13}C and ^{12}C relative abundances in *Rhodococcus* in the DNA SIP (Figure 2.8) show a large, and significant differences between the ^{13}C and ^{12}C fractions towards the heavier fractions, indicating that *Rhodococcus* spp. were incorporating carbon from isoprene into its DNA. However, when realistic levels of isoprene were added, differences in *Rhodococcus* relative abundance were small and insignificant (Figure 2.9).

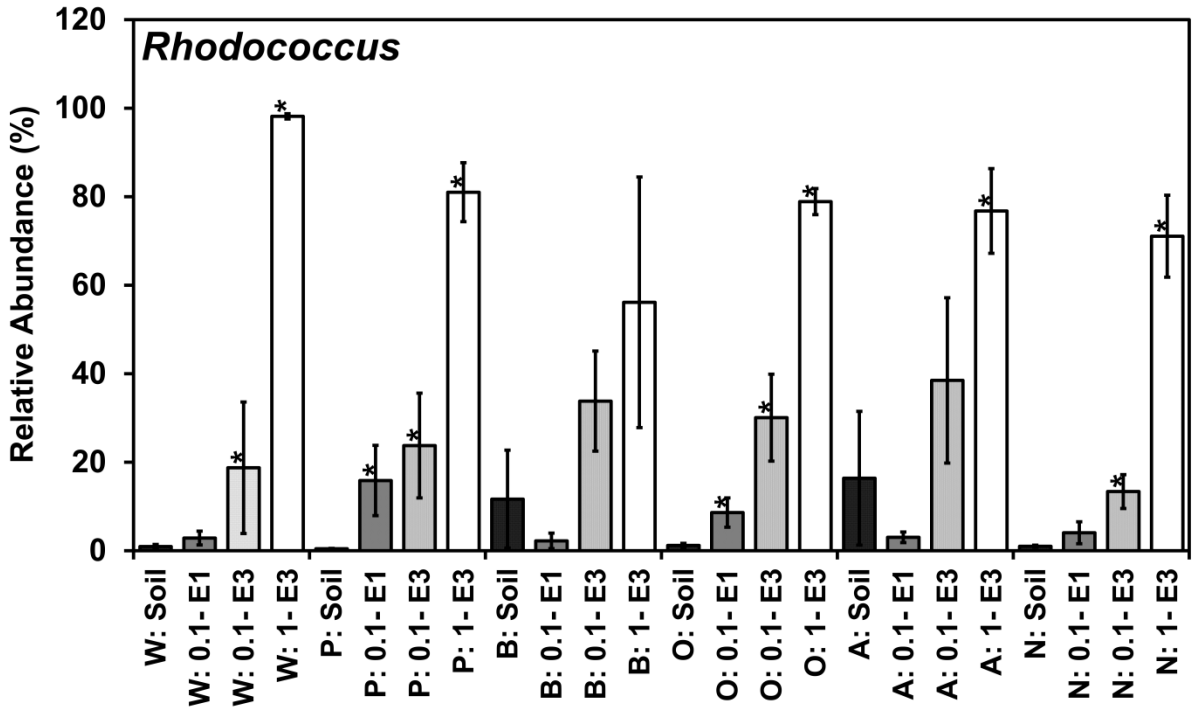


Figure 2.7 *Rhodococcus* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%; 7.2 ×10^5 ppb) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 0.78 P=0.63), was affected by concentration ($F_{8,62}$ = 34.2 P= < 0.001), and was affected by the level of enrichment ($F_{8,62}$ = 16.23 P= < 0.001), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Rhodococcus* defined by RDP classification at deepest assignment.

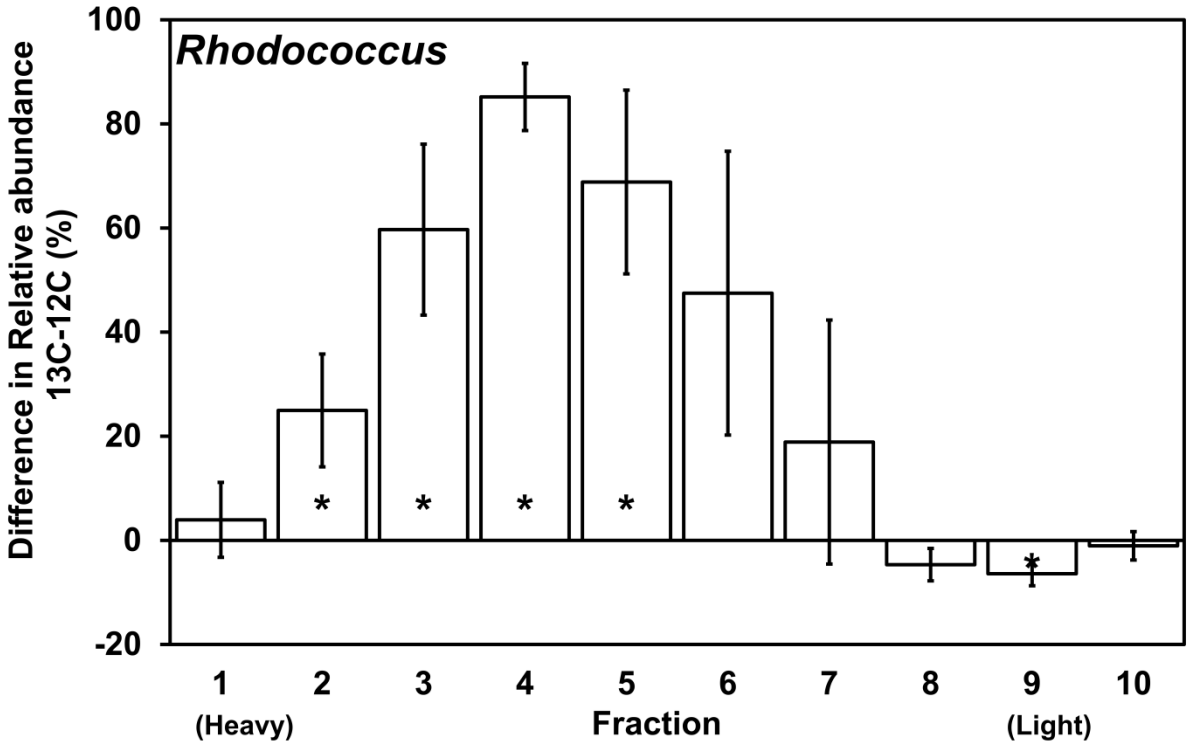


Figure 2.8 *Rhodococcus* relative abundances in ^{13}C fractions net of *Rhodococcus* relative abundances in correspondent ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by the carbon type ($F_{3,65}$ = 0.73 P=0.54), and was not affected by the density ($F_{3,65}$ = 2.52 P=0.07), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Rhodococcus* defined by RDP classification at deepest assignment.

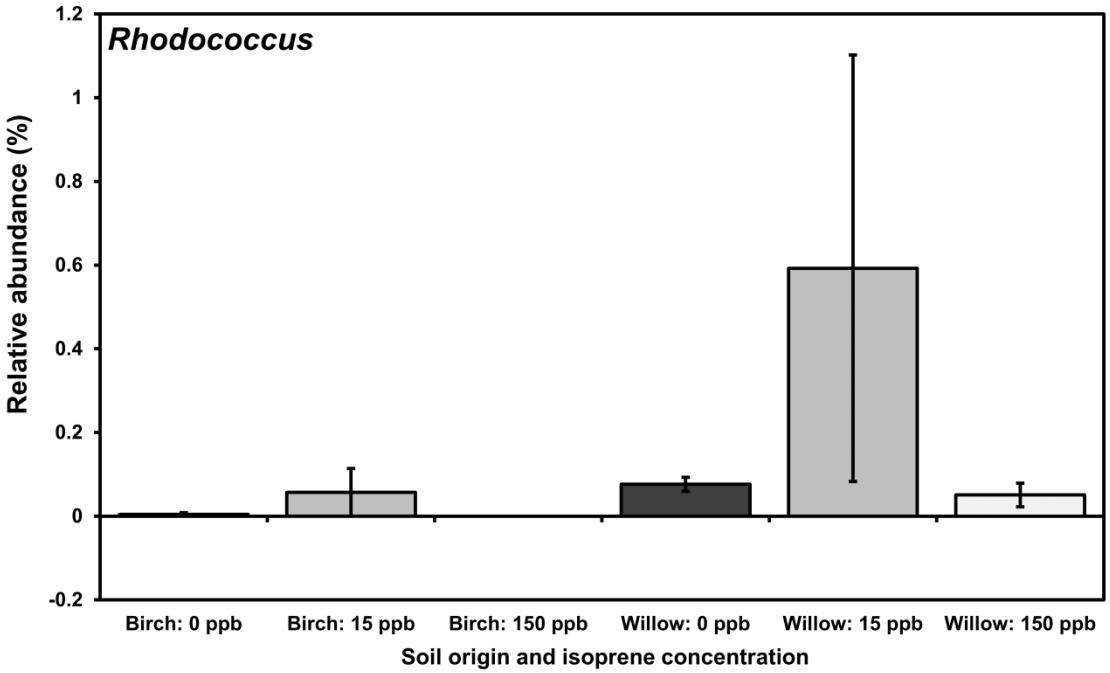


Figure 2.9 *Rhodococcus* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished 16 out of 22 days, n=3 , Error bars = SE, *Rhodococcus* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that *Rhodococcus* was not affected by the isoprene concentration ($F_{8,15}$ = 1.17 P=0.38) nor tree type ($F_{8,15}$ = 0.53 P=0.82)

OTUs classified as *Saccharibacteria* (*incertae sedis* - of uncertain placement), previously known as candidate division TM7 show increased relative abundances following sequential isoprene enrichment, with significance in 9 of 18 enrichments, including a mean of over 60% in the E3 7.2 $\times 10^5$ ppb isoprene enrichment (Figure 2.10). However, DNA SIP (Figure 2.11) demonstrated that TM7 bacteria did not incorporate carbon from isoprene into their DNA, with an increase in the proportion of ^{13}C compared to ^{12}C 16S rDNA abundance in the lighter fractions, and, no change or decreased heavier fractions. In addition to this, in the low concentration experiments (Figure 2.12), there was no noticeable change on isoprene enrichment.

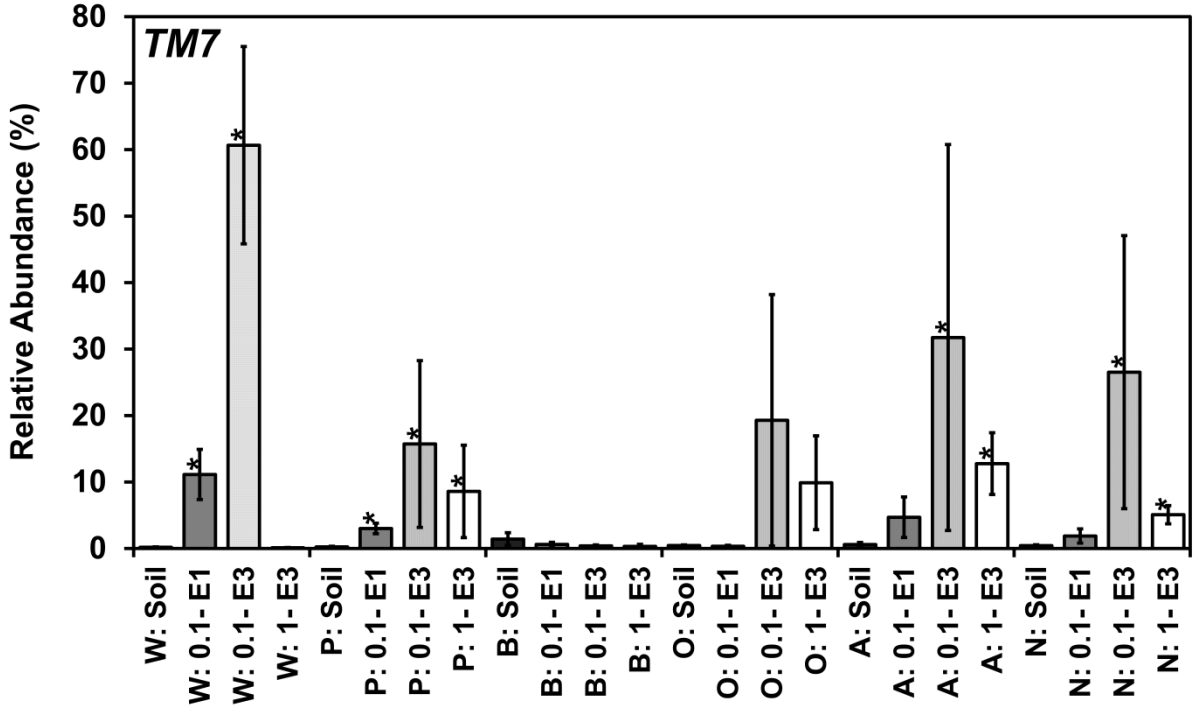


Figure 2.10 *Saccharibacteria* (aka TM7) relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 1.08 P=0.4), was affected by isoprene concentration ($F_{8,62}$ = 3.85 P= < 0.001) and was affected by the enrichment level ($F_{8,62}$ = 1.46 P=0.19), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Saccharibacteria* defined by RDP classification at deepest assignment.

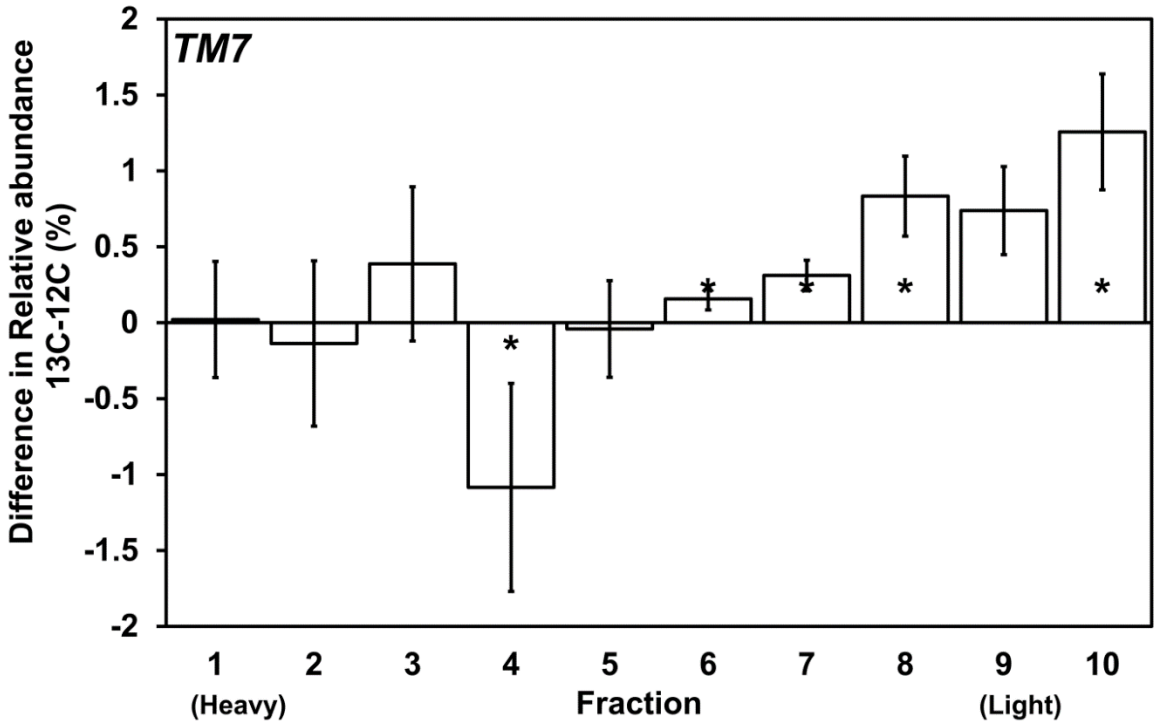


Figure 2.11 *Saccharibacteria* (aka TM7) relative abundances in ^{13}C fractions net of *Saccharibacteria* relative abundances in correspondent ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Saccharibacteria* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by the carbon type ($F_{3,65}$ = 1.35 P=0.27), and was not affected by the density ($F_{3,65}$ = 2.43 P=0.07), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Saccharibacteria* defined by RDP classification at deepest assignment.

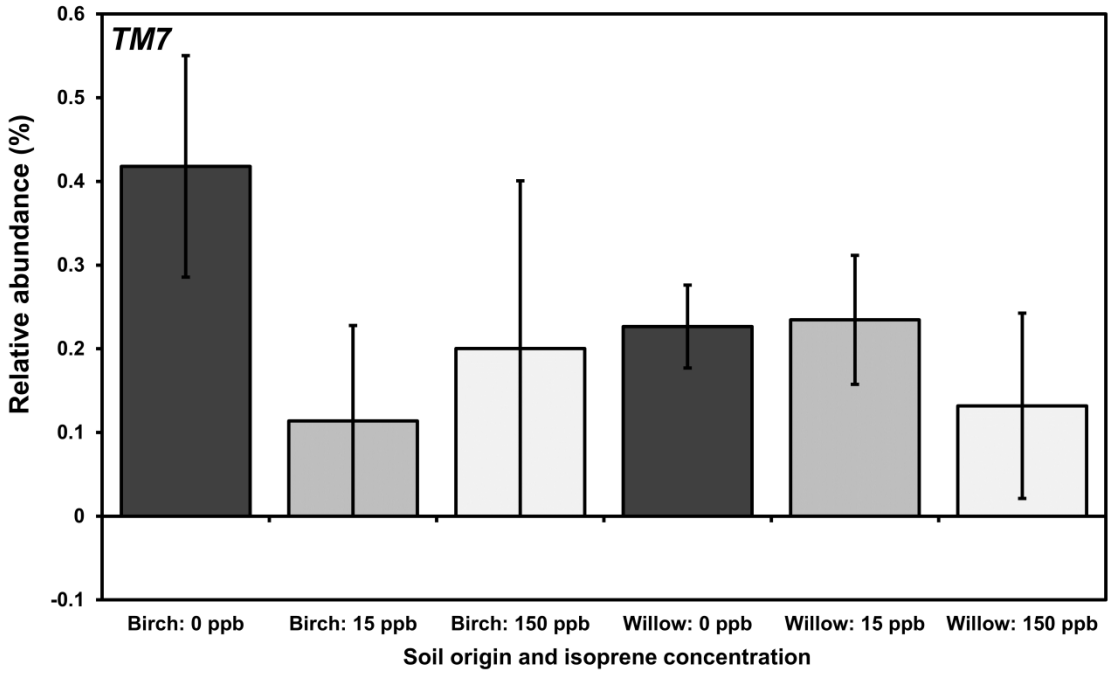


Figure 2.12 *Saccharibacteria* (aka TM7) relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Saccharibacteria* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that *abundance* was not affected by the isoprene concentration ($F_{8,15}$ = 0.76 *P*=0.64) nor tree type ($F_{8,15}$ = 0.84 *P*=0.59).

Table 2.2 Summary of OTUs showing enrichment in multiple experimental arms, with at least one significant result by sequential enrichment, DNA SIP or low concentration enrichments.

Classification	Enrichment				SIP			Low Concentration			Supplementary information Figure
	Effect of concentration	Enrichment level	$\tilde{x}_E > \tilde{x}_0$	p <0.05	Effect of density	Heavy shift	Low concentration	p <0.05	$\tilde{x}_E > \tilde{x}_0$	p <0.05	
Rhodococcus	✓	✓	16/18	11	✗	✓	✗	5	2/4	0	N/A
TM7	✓	✗	13/18	8	✗	✗	✗	(-5)	1/4	0	N/A
Aeromicrobium	✓	✓	15/18	10	✗	✗	✗	(-1)	0/4	0	(2.SI.)14,15,16
Nocardioides	✓	✗	11/18	1 (-1)	✗	✗	✗	(-1)	1/4	0	(2.SI.)17,18,19
Fluviicola	✓	✓	14/18	11	✓	✗	✓	(-1)	2/4	0	(2.SI.)20,21,22
Rhodanobacter	✗	✓	5/18	3 (-1)	✓	NP	✗	NP	0/4	(-1)	(2.SI.)23,24
Flavobacterium	✓	✗	13/18	4	✗	✗	✗	(-2)	0/4	0	(2.SI.)25,26,27
Acidovorax	✓	✓	18/18	13	✓	?	✗	0	1/4	0	(2.SI.)28,29,30
Streptomyces	✗	✓	4/18	2 (-9)	✗	✗	✗	(-1)	2/4	(-1)	(2.SI.)31,32,33
Sporolituus	✗	✗	11/18	7	✓	✗	✗	(-1)	0/4	0	(2.SI.)34,35,36
Pseudomonas	✗	✗	10/18	7	✗	?	✗	0	2/4	(-1)	(2.SI.)37,38,39
Lysobacter	✗	✗	9/18	4 (-6)	✗	✗	✗	(-1)	3/4	1	(2.SI.)40,41,42
Variovorax	✗	✗	11/18	2 (-2)	✗	?	✗	1‡(-1)	1/4	0	(2.SI.)43,44,45
Pedobacter	✓	✓	15/18	9	✗	✗	✗	0	2/4	0	(2.SI.)46,47,48
Ramlibacter	✓	✗	14/18	6	✗	NP	✗	NP	0/4		(2.SI.)49,50
Mucilaginibacter	✗	✓	6/18	2(-1)	✗	NP	✗	NP	0/4	(-1)	(2.SI.)51,52
Aeromonas	✗	✗	7/18	3	✗	NP	✗	NP	NP	NP	(2.SI.)53
Shinella	✓	✗	13/18	9	✗	✗	✗	(-1)	2/4	0	(2.SI.)54,55,56
Bosea	✓	✗	11/18	5(-3)	✓	NP	✗	NP	1/4	0	(2.SI.)56,57
Acetonema	✗	✗	11/18	3	✗	NP	✓	NP	0/4	0	(2.SI.)58,59
Polaromonas	✓	✗	13/18	3	✗	NP	✗	NP	0/4	0	(2.SI.)60,61
Caenimonas	✗	✗	12/18	3	✗	NP	✗	NP	1/4	0	(2.SI.)62,63
Acinetobacter	✗	✗	6/18	1	✓	?	✗	0	2/4	(-1)	(2.SI.)64,65,66
Methylobacterium	✗	✗	1/18	0	✓	?	✗	0	3/4	1	(2.SI.)67,68,69
Nakamurella	✓	✗	1/18	(-10)	✓	✗	✓	0	2/4	1	(2.SI.)70,71,72

Number of p values < 0.05 is by Kruskall-Wallis, with number of p values in opposite direction as negative. ‡ indicates where light fraction is significantly decreased, NP indicates not present (below the detection limit. Sorted by number of significant enrichments (all experiments). $\tilde{x}_E > \tilde{x}_0$ values are the number of samples where the isoprene-enriched relative abundance is significantly greater than the control relative abundance, out of the maximum possible. Effect of concentration shows if there was a significant interaction between concentration and abundance in the enrichment experiments, enrichment level shows if there was a significant interaction between the enrichment level and abundance in the enrichment experiments, effect of fraction shows if there was an interaction between the density and the abundance in the SIP

experiments and low concentration shows if there was a significant effect of isoprene on abundance in the low concentration experiments. Tests of interaction performed by multi-variable modelling using manyglm within the R package mvabund (details in methods, statistics in graphs).

Other groups showing multiple significant enrichments are shown in the Table 2.2. *Methylobacterium* was also included due to: i) an isolate degrading isoprene (Chapter 3), ii) it being one of the few enriched genera at low concentration (although no significant interaction was observed between *Methylobacterium* and concentration in the multi-variable model), and iii) due to a possible heavy shift in DNA density in the DNA SIP experiments. Most groups show significant enrichment on multiple enrichments at the higher isoprene concentrations (Table 2.2), with significance in DNA SIP and low concentration enrichments more rare.

Sporichthya, which are the only genus which may be significantly enriched in isoprene in the study of Gray et al. (2015), did not show any signs of enrichment in the serial enrichment (2.SI.77) experiments, and no significant change in low concentration experiments (2.SI.78).

Recent work from UEA (Crombie et al., 2016) has identified several genera of the Comamonadaceae, and *Xanthomonas, Devosia, Mycoplasma, Luteimonas*, *Mycobacterium* in leaf SIP enrichments with isoprene. In this set of experiments, several species of Comamonadaceae show signs of enrichment, although with highly sporadic presence in enrichments – including up to 1% in one enrichment, and 0.3% in one fraction of SIP, with very low starting abundances (<0.005%), however seemingly only with one replicate each time, and no significance. However, the genera *Xanthomonas* and *Mycoplasma* were not detected in the experiments in this thesis. In this chapter *Devosia* paints an interesting picture, with some OTUs (e.g. OTU 138, 97, 103) significantly enriched at least once, although at very low relative abundance (and could be by chance), and others significantly decreasing and many more having no change. Although, as a genus, *Devosia* shows a significant decrease in the heavy fraction of SIP. Taken together these studies do not provide convincing evidence of *Devosia* being a major player in isoprene-degradation in the soils tested in this thesis, but it remains a possibility. With the genera *Luteimonas*, although shown to enrich significantly a couple of times, the abundance is low, and crucially, it is shown to significantly decrease in low concentration experiments. Therefore, in soils, it is unlikely that it is involved in isoprene degradation, and if it is, it is even more unlikely to be involved at relevant concentrations.

2.4.7 OTUs of bacterial groups enriched by isoprene addition

Investigating the OTUs of groups meeting the inclusion criteria from the stable isotope probing, sequential enrichment, and low concentration experiments for OTU level change yielded a number of OTUs potentially involved in isoprene

degradation including: *Acetonema* OTU 4 (Figure 2.SI.82), *Acetonema* OTU 6 (Figure 2.SI.83), *Acidovorax* OTU 30 (Figure 2.SI.84), *Acidovorax* OTU 4 (Figure 2.SI.85), *Acidovorax* OTU 46 (Figure 2.SI.86), *Aeromicrobium* OTU 240 (Figure 2.SI.87), *Aeromonas* OTU 35 (Figure 2.SI.88), *Aeromonas* OTU 47 (Figure 2.SI.89), *Bosea* OTU 2 (Figure 2.SI.90), *Bosea* OTU 44 (Figure 2.SI.91), *Caenimonas* OTU 9 (Figure 2.SI.92), *Flavobacterium* OTU 162 (Figure 2.SI.93), *Fluviicola* OTU 190 (Figure 2.SI.94), *Fluviicola* OTU 50 (Figure 2.SI.95), *Methylobacterium* OTU l1 (Figure 2.SI.96), *Mucilaginibacter* OTU 268 (Figure 2.SI.97), *Nocardia* OTU 11 (Figure 2.SI.98), *Nocardioides* OTU 194 (Figure 2.SI.99), *Nocardioides* OTU 84 (Figure 2.SI.100), *Pedobacter* OTU 105 (Figure 2.SI.101), *Pedobacter* OTU 44 (Figure 2.SI.102), *Pedobacter* OTU 47 (Figure 2.SI.103), *Pedobacter* OTU 97 (Figure 2.SI.104), *Pelomonas* OTU l1 (Figure 2.SI.105), *Pseudomonas* OTU 27 (Figure 2.SI.106), *Pseudomonas* OTU 34 (Figure 2.SI.107), *Pseudomonas* OTU 35 (Figure 2.SI.108), *Pseudomonas* OTU 64 (Figure 2.SI.109), *Ramlibacter* OTU 19 (Figure 2.SI.140), *Rhodanobacter* OTU 24 (Figure 2.SI.141), *Rhodococcus* OTU 1473 (Figure 2.SI.142), *Rhodococcus* OTU 1908 (Figure 2.SI.143), *Rhodococcus* OTU 59 (Figure 2.SI.144), *Rhodococcus* OTU 879 (Figure 2.SI.145), *Rhodococcus* OTU s2 (Figure 2.SI.146), *Rhodococcus* OTU s3 (Figure 2.SI.147), *Rhodococcus* OTU s4 (Figure 2.SI.148), *Rhodococcus* OTU s5 (Figure 2.SI.149), *Shinella* OTU 1 (Figure 2.SI.14Figure 2.SI.150), *Sporolituus* OTU 37 (Figure 2.SI.15Figure 2.SI.153), *Sporolituus* OTU 51 (Figure 2.SI.15Figure 2.SI.154), TM7 OTU 850 (Figure 2.SI.14Figure 2.SI.151), TM7 OTU 851 (Figure 2.SI.152), *Variovorax* OTU 58 (Figure 2.SI.153), *Variovorax* OTU 62 (Figure 2.SI.154), and *Variovorax* OTU 72 (Figure 2.SI.155).

Investigating the predicted evolutionary relationships between these groups, and their respective type strains shows a wide range of diversity (Figure 2.14). The isoprene-enriched OTUs included Actinobacteria (mainly Corynebacterineae) (top left), Bacteroidetes (right), Betaproteobacteria (mainly Comamonadaceae) (top right), Gammaproteobacteria (bottom), Alphaproteobacteria (Bottom, slight right), Firmicutes (left of Gammaproterobacteria), as well as the representatives from the Candidate division TM7 (below Actinobacteria). The type strain sequences from the List of prokaryotic names with standing in nomenclature (LPSN) identified using BLASTn hits usually grouped near where they are expected to from the RDP alignment.

Figure 2.13 Evolutionary relationship between OTUs enriched by isoprene addition, and their nearest type strains, as determined from near BLASTn hits, with type strain sequences obtained from LPSN and refined iteratively. Alignment using MUSCLE (Mega 6), Relationship predicted using Jukes and Cantors model of Neighbour Joining with 1000 Bootstrap tests (values inset in grey) (Jukes and Cantor, 1969; Saitou and Nei, 1987; Felsenstein, 1985). OTU names from RDP classifier deepest assignment, "OTU l" and "OTU s" OTUs are from low concentration experiments and SIP experiments respectively, Evolutionary analysis performed in MEGA 6 (Tamura et al., 2013). Scale is nucleotide substitutions per site.

The OTU centroids derived from the SIP experiments, which were generated independently, group very closely to some of the groups from the sequential enrichments, including the SIP OTU *Rhodococcus* OTU s2 clustering closely with *Rhodococcus* OTU 1908, *Rhodococcus* OTU s3 grouping closely with

Rhodococcus OTU 59, and *Rhodococcus* OTU s4 very closely to *Rhodococcus* OTU 1473.

TM7, being a candidate division, has no type strains, however the OTU centroid sequences were similar to genome sequenced TM7 sequences. TM7 OTU 850 had 98% identity to “GW2011_GWC2_44_17”, from Brown et al. (2015), and TM7 OTU 851 had 97% identity to “RAAC3_TM7_1” from Kantor et al. (2013).

2.5 Discussion

2.5.1 In situ rates of isoprene degradation in soils were near the lower measurements, suggesting overestimation of the size of the global soil sink.

Rate testing for in situ soil isoprene degradation throughout the year yielded no significant differences between the days tested across different seasons, with an average consumption rate of 4.77 nmol m^{-2} h^{-1}. This is notably similar to the in situ rates reported by Gray et al. (2014), with an average of 2.0 nmol m^{-2} h^{-1} with a max of 7.4 nmol m^{-2} h^{-1}. However, this is much lower than the rate of 1428 nmol m^{-2} h^{-1}reported by Cleveland and Yavitt (1997). Aside from the location and sampling method, the starting isoprene concentration is probably the key factor that led to such between study differences. Cleveland and Yavitt (1997) added 385 ppb, Gray added no isoprene, and looked at the natural flux, and in this study, only 15 ppb was added. As Pegoraro (2005) demonstrated, there are clear, positive, relationships between isoprene concentration and isoprene-degradation rates, with the exact effect depending on other factors (e.g. temperature and soil water activity). Cleveland and Yavitt (1997) likely overestimated the natural isoprene flux by far exceeding realistic isoprene concentrations, and so their estimation of the global soil isoprene sink of ~20.4 Tg yr^{-1}, is likely significantly too high. Cleveland and Yavitt (1997) suggested as much themselves, putting it forward as an upper bound, noting that actual levels of isoprene would be nearer to 3 ppb. It is probable that the isoprene consumption rates determined by Gray et al. (2015), which used unadjusted concentrations, are more realistic reflection of true rate. This finding is supported by the data in this thesis (Fig 7.2 $\times 10^6$) and extrapolations from

Pegaroro et al. (2005), which near zero degradation when you trend to atmospherically relevant isoprene concentrations (Chapter 1). Based on the mean value determined in this thesis (4.77 nmol m^{-2} h^{-1}) and that obtained by Gray et al. (2015) (mean of 2.0 nmol m^{-2} h^{-1}, max of 7.4 nmol m^{-2} h^{-1}), the mass of atmospheric isoprene degraded by soil would be in the region of 0.03 Tg yr^{-1} (Here: 0.068 Gray mean: 0.029, Gray max: 0.11 Tg yr^{-1}). This means that previous estimates of isoprene consumption in soil would have given an overestimate of soil as a sink, and soils may be near irrelevant as atmospheric isoprene sinks. It is likely that the effect of soil degradation in those models will require substantial revisions downwards; for example, the figure of 20.4 Tg isoprene degraded per year by soils, could be over 100 times the actual value (although as half the global flux determined from that work was from one forest soil type with high variance; this should not be too much of a surprise), although more research is needed to clarify this.

However, firstly this is based on two small scale experiments, and so not representative at a global scale. Many more experiments would be required for an accurate global rate measurement; and secondly, as increasing the isoprene concentration does increase the rate by a large amount, there are likely to be some other factors involved as the soils are clearly primed for isoprene concentrations far exceeding those which they would be exposed to from the atmosphere.

2.5.2 No difference in degradation rates is observed between soils from beneath or remote from isoprene-producing trees.

Using ex situ laboratory incubations with isoprene concentrations of 7.2×10^5 ppb and 7.2×10^6 ppb, soils taken from under the canopies of isoprene-producing trees have the same rates of isoprene degradation as those from areas with no trees, or non-isoprene-producing trees. However, isoprene degradation was very rapid, six times the rate found in microcosms of estuary water (Acuña Alvarez et al., 2009), and ten times the speed of lake water (not shown), suggesting the presence of microbiota able to quickly take advantage of the presence of isoprene. Transferring soil to a minimal medium and adding a high concentration of isoprene would be expected to cause a lag phase, the presence of which would effectively reduce the difference in degradation rate, so that the total rate difference is likely greater. This lack of significant difference in degradation rates between samples from near, and away from what we believed to be the key production sources, combined with the expectation that (discounting wind) the atmospheric isoprene concentration should follow an inverse squared effect from the source, along with rapid loss to atmospheric reactions generates the question: if the isoprene-degrading ability of bacteria isn't significantly driven by nearby tree type, what is driving the ability to degrade isoprene? Some soil dwelling bacteria, particularly *Bacillus* species (Julsing et al., 2007), produce isoprene, which raises the possibility of an in-soil microbial isoprene cycle.

However, the stands in this experiment were small, and the isoprene concentration used in the incubations was high, so it is possible that in large

monoculture forests, with atmospheric levels of isoprene, there may be noticeable differences.

2.5.3 Isoprene-degrading bacteria are numerous, varied, and possibly specialised for different isoprene concentrations

A four-pronged experimental design was deployed to investigate the isoprene-degrading microbiota: (i) sequential enrichment in minimal media with high concentrations of isoprene was used to enrich isoprene degraders, (ii) plating and isolation was used to cultivate isoprene degraders from these enrichments, (iii) stable isotope probing with ^{13}C-isoprene was used to specifically detect bacteria that incorporated isoprene into their biomass, and (iv) media free low isoprene concentration experiments were used to try and draw apart those bacteria that were likely to degrade isoprene at near atmospheric concentrations and so, were more likely to be involved in isoprene degradation in situ.

The sequential enrichments with high concentrations of isoprene show a clear effect of isoprene on bacterial community composition, driven mainly by increasing the extent of enrichment, and secondly by isoprene concentration (Figure 2.3). Overlapping clustering of the third (E3) enrichments with both isoprene concentrations suggests common elements, likely one or more bacterial groups being enriched across the samples. Additionally, the highest level enrichments clustering more tightly than other enrichments suggests a further narrowing of the community structure there than at the lower high level enrichment. This convergence suggests a clear, concentration mediated, enrichment level specific enrichment, resulting in a somewhat replicable bacterial community composition. Thus, this sequential enrichment experiment

has sufficient power and resolution to identify species involved in isoprene degradation, which will be largely starting sample independent. As seen in Table 2.2 and Figure 2.7, it is likely that a large part of this effect is due to enrichment of *Rhodococcus* species.

The detrended correspondence analysis from the DNA SIP experiment shows that upon isoprene enrichment, there is a clear effect distinguishing the ^{13}C fractions from the ^{12}C fractions with increasing fraction density, most easily visible in Figure 2.6, combined with a slight pull in the opposite direction for the lighter fractions (due to absence of heavy fraction community members), and that the initial communities cluster more with the ^{12}C and lighter ^{13}C fractions – demonstrating that the addition of ^{13}C isoprene increases the density of a portion of the DNA, and therefore that isoprene is being directly incorporated by some species. In Figure 2.5, large community changes in some of the heavy fractions with both ^{12}C and ^{13}C isoprene are visible, possibly representing community change of GC rich bacteria, contamination in these fractions upon tube straightening, or possibly just the stochastic nature of community assembly.

It is worth noting, that when analysing the DNA SIP data, the graphs should be interpreted with the following in mind: (i) the total DNA concentration is likely higher in the light than heavy fractions, (ii) different species of the same genus could act in different ways, (iii) data is grouped by fraction rather than density in the graphs, as a proxy for density, and points were coloured by density as approximated by the relative index of the fractions for the ordination plots.

With the low concentration experiments, isoprene has a visually noticeable effect, negatively on the NMDS2 axis, however this was insignificant compared to the no isoprene additions. Although this raises difficulties in terms of determining which bacteria are enriched – it also demonstrates that the isoprene concentrations provided to the soil are close to what they are used to; somewhat validating that soil isoprene concentrations are in the rough region of the experiment (and likely <15 ppb). The clear differences in orientation of communities from willow and birch soils, with clear separation on the NMDS1 axis, and no obvious signs of convergence on enrichment is quite unlike the high concentrationenrichments (7.2×10^6 ppb) which became more similar when enriched, and is probably due to other factors (e.g. pH or nutrient differences), and suggests that the convergence seen in higher isoprene enrichments may be due to those experimental designs causing specific enrichment of a small subset of a more diverse isoprene-degrading community. There was also a clear difference in size of the effect between the willow and the birch samples, with the birch samples showing greater movement on isoprene enrichment (i.e. a larger difference in the position of points on ordination with different concentrations of isoprene), demonstrating that isoprene had a greater effect on the birch microbial community than the willow microbial community composition. This could be due to the willow sample being more used to isoprene, or possibly just an effect of the 2d visualisation.

Isoprene degraders were initially grouped and screened at the RDP “Fix Rank” classification level, and if significant in multiple enrichments, then scrutinised at the OTU level. As demonstrated by the wide variety of bacterial classes present (Figure 2.13), the strains which are enriched by isoprene are diverse, whether

directly degrading isoprene, or not, in some measure the effects of isoprene are not localised to a small group of bacteria.

The serial enrichments show an abundance of *Rhodococcus*, increasing through the sequential levels of enrichment, and especially in later and higher isoprene concentration enrichments. In addition, the dominance of *Rhodococcus* in the DNA SIP heavy fractions, with a significant decrease in the light fractions, clearly demonstrate the incorporation of carbon derived from heavy isoprene into *Rhodococcus* DNA. This affirms what we already knew about the genus *Rhodococcus* containing isoprene degraders (van Hylckama Vlieg et al., 2000). Additionally, numerous species of *Rhodococcus* were isolated with isoprene as a sole carbon source in attempts to build a collection of isoprene-degrading organisms as part of this project (Chapter 3).

However the changes in abundance on enrichment with low concentrations of isoprene were unconvincing, with no significant dominance of enriched relative abundances, suggesting that *Rhodococcus* may not play a role in the uptake of isoprene at low concentrations, although clearly primed to take advantage of high levels of isoprene.

Despite Gray et al. (2015) having issues due to the lack of significance shown (assuming present) and unclear inclusion criteria, the absence of *Rhodococcus* in their OTU level analysis is conspicuous, and backs up the idea that *Rhodococcus* is not important in isoprene degradation at low concentrations.

Rhodococcus abundance changes with differing concentrations of isoprene were significant ($F_{8,62}$ = 34.2 P= < 0.001) in the enrichment experiments, and *Rhodococcus* abundance changes with differing concentrations of isoprene

were not significant ($F_{8,15}$ = 1.17 P=0.38) in the low concentration isoprene experiments. *Rhodococcus* abundances at different densities were not significant ($F_{3,65}$ = 2.52 P=0.07), and with carbon types were not significant ($F_{3,65}$ = 0.73 P=0.54) in the SIP isoprene experiments.

At the OTU level, eight *Rhodococcus* OTUs seemed to be involved in isoprene degradation, four from enrichments (OTU 1908 (Figure 2.SI.144), 1473 (Figure 2.SI.143), 59 (Figure 2.SI.145), 879 (Figure 2.SI.146), and four from stable isotope probing (OTU s2 (Figure 2.SI.149), s4 (Figure 2.SI.151), s3 (Figure 2.SI.150), s1 (Figure 2.SI.148)). *Rhodococcus* 1908 and s2 are very similar and are most closely related to the type strain *Rhodococcus globerulus* (Figure 2.12). The SIP and low concentration experiments were performed separately (~ 2 years apart), and bioinformatically analysed separately; so this similarity suggests that the results are comparable and robust. *Rhodococcus globerulus* is one of the better known isoprene degraders. *Rhodococcus* AD45, the strain in which the isoprene degradation pathway was discovered, is likely a member of the species *Rhodococcus globerulus* (Johan et al., 1998), and several *R. globerulus* strains have been shown to degrade isoprene (Johnston., 2014; El Khawand, 2014; van Hylckama Vlieg et al., 2000), so it is predictable that it is again shown to be involved in isoprene degradation. *Rhodococcus* OTU 59 and s3 are also closely related to each other and most similar in partial 16S rRNA gene sequence to *Rhodococcus wratislaviensis*; again a previously known isoprene degrading species (Crombie et al., 2015). *Rhodococcus* OTU s4 and *Rhodococcus* OTU 1473, likewise, are closely related to *Rhodococcus erythropolis*, which is another known isoprene degrader, and similar to *R.*

globerulus and AD45 (Crombie et al., 2015). *Rhodococcus* OTU 879 is likely a strain of *Rhodococcus opacus*, also known to be an isoprene degrader (Crombie et al., 2015). So whether or not *Rhodococcus* is involved in isoprene degradation in realistic atmospheric concentrations, the genus contains a number of species which directly degrade isoprene and are rapidly and significantly enriched from environmental samples, seemingly regardless of where the sample is derived from (including soils, fresh water and seawater (Johnston, 2014); although this does raise the additional question of whether *Rhodococcus* is an isoprene degrading lab weed). If *Rhodococcus* species are not involved in isoprene degradation with environmentally relevant concentrations, then this suggests that something else might be, and that there may be different bacteria primed for different concentrations of isoprene. Alternatively (but less likely), this difference in enrichment at different isoprene levels could suggest that in nature isoprene is not the main target of degradation by *Rhodococcus* spp., and the isoprene degrading capacity studied so far is a byproduct of a different reaction. The isoprene synthase active site is not highly specific, some isoprene-degrading bacteria co-metabolise trichloroethene, dichloroethenes, vinyl chloride, and Dichloroepoxyethane, potentially causing damage to the organism in the process (Ensign et al., 1992; Johan et al., 1998). However, it is also possible that, in the case of isoprene production in the soil, micro-environments rich in isoprene could lead to benefits for low affinity isoprene degraders.

Candidate division TM7 (also known as *Saccharibacteria*) is enriched by isoprene in the sequential enrichments; in one enrichment reaching a mean of over 60%. However, these increases in abundance were highly stochastic,

generating large error bars and absences, and primarily present in the lower level (7.2×10^5 ppb) sequential enrichments. However; the DNA SIP clearly evidences that TM7 does not incorporate isoprene into its cellular macromolecules at this concentration, with multiple increases in the lighter fractions relative abundance (net of control), many of which were significant, and significant decreases in some of the heavier fractions is shown - demonstrating that although TM7 is enriched, it is only enriched in the ^{12}C arm of the experiment. Although this could be due to a peculiarity of TM7 not being able to use ^{13}C isoprene (in a similar way to some *Methanotrophs* having lower affinity for ^{13}C methane (Bull et al., 2000)), it is much more likely that TM7 is not directly using isoprene, and the increased prevalence is due to indirect enrichment (TM7 is clearly gaining something from the isoprene addition). This is backed up by what we know about some TM7 species requiring other organisms to survive with the limited cultivation requiring presence of other organisms, and the lack of some TM7 species to synthesise key molecules, e.g. amino acids, and to be essentially parasitic (He et al., 2015). Additionally, TM7 is not enriched in the low concentration experiments, suggesting irrelevance in real world isoprene removal, and the indirect effect of enrichment being unlikely in situ. Interestingly however, TM7 was shown to be present on 4/9 leaves from isoprene-producing tobacco, and 0/9 of non-producing tobacco (Chapter 4), although the significance of this is still unclear. TM7 abundance changes with differing concentrations of isoprene was significant ($F_{8,62} = 3.85$ $P = < 0.001$) in the enrichment experiments, and TM7 abundance changes with differing concentrations of isoprene were not significant ($F_{8,15} = 0.76$ $P=0.64$) in the low concentration isoprene experiments. TM7 abundances at different densities

were not significant ($F_{3,65}$ = 2.43 P=0.07), and with carbon types were not significant ($F_{3,65}$ = 1.35 P=0.27) in the SIP isoprene experiments.

At the OTU level, TM7 falls into two groups, TM7 OTU 850 (Figure 2.SI.155) has high similarity to one of the TM7s genome sequenced by Brown et al. (2015), however little is known about it yet. TM7 OTU 851 (Figure 2.SI.156) was similar to the sediment derived TM7 sequenced by Kantor et al. (2013). This TM7 lacks the pentose phosphate pathway, most of the EMP pathway and no enolase, or method of converting pyruvate to acetyl-CoA, few genes involved in nucleotide biosynthesis, as well as some other usual metabolic pathways being missing. This suggests that TM7 may be reliant on metabolites from other bacteria to thrive, and possibly that these bacteria are being enriched. Linear regression was performed between TM7 relative abundance and the relative abundance of every other fix rank “genus” for each of these experiments, however relationships were only valid within experiments; meaning that any relationship to other bacteria is either unlikely, or complex without direct correlation, or with second species redundancy (i.e. multiple species which it can grow in the presence of, where only one needs to be present). Similarly, this lack of clear dependence is supported by co-occurrence analysis; bacteria that co-occurred with TM7 were different in each experiment (supported further by the fact that TM7 reaches 89.8% in one enrichment, with no more than 3% relative abundance of any other genus level group; so any indirect effect would have to support at least 30 TM7 for each bacteria creating the effect). Despite the uncertainty of exactly what TM7 is doing to become enriched upon isoprene addition, certainly at high levels it is continuously enriched. This could be due to its growth depending on another bacterium

enriched by isoprene, or a factor of the experimental design/conditions, and further investigation would be required to try and explain its enrichment. The TM7 issue also demonstrates the reliability of the multi-faceted experimental approach; without SIP, we may have thought of it as an isoprene degrader; and without low concentration experiments, we may have further thought that it was involved in in situ isoprene degradation. Interestingly, TM7 has been shown to degrade Benzene, and Toluene (somewhat similar compounds) through SIP (Luo et al., 2009, Xie et al., 2010) (Chapter 3 explores TM7 in greater depth).

Aeromicrobium species increase with sequential enrichments, even reaching means of over 50% in two cases (Figure 2.SI.14), however *Aeromicrobium* was not shown to be directly involved in isoprene degradation by SIP, and showed no signs of being involved in isoprene degradation in the low concentration experiments. Additionally, an *Aeromicrobium* species (61[ba]) was isolated using isoprene as a sole carbon source, despite *Aeromicrobium* species often requiring precise nutritional profiles in the growth media to cultivate (Hernandez-Eugenio et al., 2015); however it failed to degrade isoprene in pure culture (Chapter 3), adding evidence that whatever factor is causing *Aeromicrobium* growth in the sequential enrichments is likely not directly related to either isoprene, and possibly not due to other bacteria using isoprene (as *Aeromicrobium* was able to be isolated without use of other bacterial input), and more likely due to the experimental conditions (e.g. the minimal media, temperate). The significant increase of *Aeromicrobium* in the enrichment was mainly contributed to by one OTU, OTU 240, which reaches a mean relative abundance of ~50% in the third “0.1” level isoprene enrichment (7.2×10^5 ppb) for Poplar and No Trees soil, and was similar to the *Aeromicrobium erythraeum*

type strain, and *Aeromicrobium* does not specifically appear in the literature as an isoprene degrader (although an *Aeromicrobium* does appear in Sample 6 ^{13}C, and slightly more in ^{12}C in El Khawand (2014), and also *Aeromicrobium* is present at low level in the "Indonesia Water" isoprene enrichment by Acuña-Alvarez et al., (2009), however no initial concentrations or significance was stated for these). Although it is probably not a direct isoprene degrader, *Aeromicrobium* does seem to become indirectly enriched, sometimes to over half of the population, meaning it is likely involved somewhere in subsequent steps in the metabolism of carbon from isoprene. *Aeromicrobium* abundance changes with differing concentrations of isoprene was significant ($F_{8,62}$ = 3.19 P=0.01) in the enrichment experiments, and *Aeromicrobium* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 1.43 P=0.26) in the low concentration isoprene experiments. *Aeromicrobium* abundances at different densities were not significant ($F_{3,65}$ = 2.18 P=0.1), and with carbon types were not significant ($F_{3,65}$ = 1.12 P=0.35) in the SIP isoprene experiments.

Mycobacterium species, some of which have previously been shown to be involved in isoprene degradation, and has been isolated an tested as a degrader (El Khawand, 2014; Johnston, 2014), did not show any signs of isoprene degradation, and in fact shown significant suppression and light shifts (including decreases in relative abundance through enrichments which were statistically significant in eight samples, a significant decrease in the heavy fractions, and no change in relative abundance when incubated with environmentally relavent concentrations of isoprene) (Figure 2.SI.74, 2.SI.75,

2.SI.76), suggesting that isoprene degrading strains are not ubiquitous in these environments tested.

However, *Nocardia*, identified previously as isoprene degraders (Hou et al., 1981), showed enrichment in many of the serial enrichments (Figure 2.SI. 77, 78, 79), mainly represented by OTU 11 (which does seem to degrade isoprene). However, *Nocardia* had indeterminate SIP results and low concentration experiments, with a significant decrease in birch soil exposed to 150 ppb isoprene, potentially making their environmental relevance doubtful. The results from Johnston (2014), are contradictory, suggesting a *Nocardia* increase in the light fraction of ^{13}C SIP to >11%, with absence in the heavy fraction, although with only low presence in the ^{12}C fractions. In summary, although an isoprene degrading Nocardia isolate exists, the isoprene degrading species of this genus are unlikely to be widespread - as species which degrade isoprene are not consistently present in experiments, and species which do not are often present.

Nocardioides seems to be enriched in 7.2 $\times 10^5$ ppb isoprene sequential enrichments, and seems to decrease in abundance with higher isoprene additions, and some of the later 7.2 $\times 10^5$ ppb isoprene additions. This could be due to it being outcompeted, however the lack of trend in low-concentration enrichments, and the lack of incorporation of ^{13}C into DNA in the DNA SIP experiments means that its direct involvement in isoprene degradation in the environment is unlikely. On another note, *Nocardioides* was slightly enriched in heavy compared to light fractions of SIP in El Khawand (2014), however no significance was stated, and also, a *Nocardioides* species was isolated in this project (bA1a), on minimal media plates with isoprene (Chapter 3); so there is

good circumstantial evidence of involvement. *Nocarioides* OTUs included two similar to *N*. *caeni* (OTU 194; OTU 84), which follow the Genus' trend. *Nocardioides* abundance changes with differing concentrations of isoprene was significant ($F_{8,62}$ = 3.19 P=0.04) in the enrichment experiments, and *Nocardioides* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 0.3 P=0.96) in the low concentration isoprene experiments. *Nocardioides* abundances at different densities were not significant ($F_{3,65}$ = 1.49 P=0.23), and with carbon types were not significant ($F_{3,65}$ = 0.89 P=0.45) in the SIP isoprene experiments.

Fluviicola showed frequent, often modest but occasionally large, significant enrichment through serial enrichment of isoprene, suggesting, along with non-significant increases in low concentration enrichments of Willow soil, and a sinusoidal DNA SIP plot (forming a wave of multiple peaks and troughs along the density gradient – which on further investigation was caused by two different (but dominant) *Fluviicola* OTUs, with one being responsible for the heavy shift). This suggests possible involvement of *Fluviicola* in isoprene degradation. I would suggest that although the SIP and low concentration results do not have statistical significance, in combination with the consistent enrichment in the serial enrichments, some *Fluviicola* species are likely isoprene degraders, and may be relevant at real world levels of isoprene exposure. *Fluviicola* OTU 190 and 50 were significantly enriched in the serial enrichments, and were both similar to F. *taffensis*. There is no indication of either F. *taffensis,* nor any member of the *Fluviicola* genus, degrading isoprene in the literature, making it a candidature for further investigation. *Fluviicola*

abundance changes with differing concentrations of isoprene was significant ($F_{8,62}$ = 3.81 P= < 0.001) in the enrichment experiments, and *Fluviicola* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 0.01 P=1) in the low concentration isoprene experiments. *Fluviicola* abundances at different densities were significant ($F_{3,65}$ = 4.4 P=0.01), and with carbon types were significant ($F_{3,65}$ = 0.75 P=0.53) in the SIP isoprene experiments.

The sporadic occurrence of *Rhodanobacter* in serial enrichments combined with absence in the DNA SIP experiments and no enrichment in the low concentration experiments suggest that the significant enrichments shown in some of the serial enrichment were not directly driven by the presence of isoprene. *Rhodanobacter* OTU 24 was responsible for the changes noted (Figure 2.SI.142), and was identified as being similar to R *thiooxidans*, and like the genus as a whole, is unlikely to be involved in isoprene degradation. *Rhodanobacter* abundance changes with differing concentrations of isoprene were not significant ($F_{8,62}$ = 0.32 P=0.96) in the enrichment experiments, and *Rhodanobacter* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 1.3 P=0.32) in the low concentration isoprene experiments. *Rhodanobacter* abundances at different densities were significant ($F_{3,65}$ = 2.74 P=0.05), and with carbon types were significant ($F_{3,65}$ = 0.39 P=0.77) in the SIP isoprene experiments.

Flavobacterium showing consistent, and occasionally significant enrichment across sequential enrichments is, in a manner similar to TM7, probably unlikely to be directly driven by the presence of isoprene due to the DNA SIP experiments clearly demonstrating a decrease in the heavier fractions alongside a possible increase in the lighter fractions. This suggests that *Flavobacterium* enrichment is indirect, and the lack of enrichment in low concentrations of isoprene, suggests that it is environmentally irrelevant in the degradation of isoprene. Although it is interesting to note that some *Flavobacterium* species have a Glutathione S-transferase, which may be similar to that involved in isoprene degradation (van Hylckama Vlieg et al., 1999; Xun et al., 1992), although this is of limited applicability without further information. *Flavobacterium* are also present in the work done by El Khawand (2014), with presence in the initial time periods at low abundance, followed by variable enrichment or suppression (although no significance stated), likewise they are present in Johnston (2014) as a dominant member (1/3) of Hythe samples, going down to a few percent in the light isoprene, and are absent in the heavy isoprene samples (although with no replication or statistics). Considering the lack of enrichment, lack of ^{13}C incorporation, and the high abundance in which they can be found in water, it is possible that *Flavobacterium* are just better suited to the minimal media environment i.e. successful oligotrophs that scavange trace carbon sources, and are able to thrive in a liquid environment. *Flavobacterium* abundance changes with differing concentrations of isoprene was significant ($F_{8,62}$ = 2.14 P=0.05) in the enrichment experiments, and *Flavobacterium* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 0.56 P=0.8) in the low concentration isoprene

experiments. *Flavobacterium* abundances at different densities were not significant ($F_{3,65}$ = 0.66 P=0.59), and with carbon types were not significant ($F_{3,65}$ = 0.45 P=0.72) in the SIP isoprene experiments.

The interaction between *Acidovorax* species and isoprene is unclear. The sequential enrichment shows a ubiquitously higher mean upon enrichment, however with no clear picture from DNA SIP or low concentrations (and no more clarity at the OTU level (with different Acidovorax OTUs seemingly enriched in different enrichments)), it is hard to tell if the enrichment is directly or indirectly driven by isoprene, however being enriched significantly 13 times, and markedly increasing from the first to the third 7.2 $\times 10^5$ ppb isoprene addition enrichments, it is likely that there is some effect of isoprene driving the structure, diversity and abundance of OTUs in the population. In addition to this, although it may represent different *Acidovorax* species, a isoprene DNA SIP experiment with marine bacteria shows a visually greater relative abundance of *Acidovorax* in the light ^{12}C, compared to heavy ^{13}C fractions, although further complicated by an absence in the second ^{12}C duplicate, and an absence of statistical analysis (Johnston., 2014). This, in some way supports the notion that the enrichment of different *Acidovroax* species by isoprene is replicable, affecting a number of *Acidovorax* species, and that this enrichment is indirect. *Acidovorax* abundance changes with differing concentrations of isoprene was significant ($F_{8,62}$ = 9.26 P= < 0.001) in the enrichment experiments, and *Acidovorax* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 0.01 P=1) in the low concentration isoprene experiments. *Acidovorax* abundances at different densities were significant ($F_{3,65}$ = 3.7

P=0.02), and with carbon types were significant ($F_{3,65}$ = 0.59 *P*=0.63) in the SIP isoprene experiments.

Streptomyces, despite increasing in abundance in every first level enrichment with 0.1 ml isoprene headspace (7.2 ×10^5 ppb), with significance twice, is unlikely to contain isoprene degrading species, as by the third enrichment of both isoprene levels *Streptomyces* is reduced to little or none of the relative abundance, and is significantly decreased in the heavy fractions of DNA SIP and possibly increased in the light fractions, and has no change and one significant decrease in the low concentration experiments. It should be noted however that *Streptomyces* is present at 1.5 % in Acuña-Alvarez's (2009) High isoprene Etang De Berre enrichment (although without initial abundances, replication or stats). However, it is likely that the enrichment of *Streptomyces* in the sequential enrichment (and the other studies) was likely an artefact of early experimental conditions (e.g. presence of surfaces, or liquid media environment). *Streptomyces* abundance changes with differing concentrations of isoprene were not significant ($F_{8,62}$ = 0.52 *P*=0.84) in the enrichment experiments, and *Streptomyces* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 1.27 *P*=0.33) in the low concentration isoprene experiments. *Streptomyces* abundances at different densities were not significant ($F_{3,65}$ = 1.1 *P*=0.36), and with carbon types were not significant ($F_{3,65}$ = 0.71 *P*=0.55) in the SIP isoprene experiments.

Sporolituus and *Sporolituus* OTUs, despite going from next to no abundance in the soil to low presence (Most means 0.5-1.5%, with one over 7%) in most enrichments for most soil types is unlikely to be involved directly in isoprene degradation, showing a significant decrease in one of the heavier fractions, and no change in the low concentration experiments, suggesting any effect of isoprene is indirect. *Sporolituus* abundance changes with differing concentrations of isoprene were not significant ($F_{8,62}$ = 1.03 P=0.43) in the enrichment experiments, and *Sporolituus* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 0.01 P=1) in the low concentration isoprene experiments. *Sporolituus* abundances at different densities were significant ($F_{3,65}$ = 3.65 P=0.02), and with carbon types were significant ($F_{3,65}$ = 1.02 P=0.39) in the SIP isoprene experiments.

The effect of isoprene on the relative abundance of *Pseudomonas* is debatable, in sequential enrichments showing enrichment on many of the 0.1 ml isoprene headspace (7.2 ×10^5 ppb) enrichments, but no presence in the 1.0 ml isoprene (7.2 ×10^6 ppb) enrichments. The DNA SIP experiments are inconclusive, with a unclear (and sinusoidal) pattern, at the OTU level, however, it seems that at least one OTU is enriched but is again probably not directly involved, and at least one OTU may be – although with no significance it is hard to be sure. Combined with the low concentration experiments being equally unclear, it is hard to tell if *Pseudomonas* is directly involved in isoprene degradation from these experiments. However, two different species of *Psuedomonas* were isolated on isoprene as a sole obvious carbon source (3.88, P1) as part of this project (Chapter 3). Additionally, *Pseudomonas* was shown to degrade

isoprene at very high concentrations in Srivastva et al., 2015, and Acuña-Alvarez et al., 2009 noted *Pseudomonas* in two isoprene enrichments, and noted *Pseudomonas* as a dominant last stage DGGE band (although lacking change, replication and statistics). *Pseudomonas* is also known to degrade other small hydrocarbons, including phenol and toluene, and has monooxygenases similar in sequence to that of isoprene monooxygenase (Leahy et al., 2003; Crombie and Murrell, 2014). *Pseudomonas* clearly has isoprene degrading species, although it's enrichment may be condition dependant. *Pseudomonas* abundance changes with differing concentrations of isoprene were not significant ($F_{8,62}$ = 1.92 P=0.07) in the enrichment experiments, and *Pseudomonas* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 0.5 P=0.84) in the low concentration isoprene experiments. *Pseudomonas* abundances at different densities were not significant ($F_{3,65}$ = 0.4 P=0.76), and with carbon types were not significant ($F_{3,65}$ = 0.81 P=0.49) in the SIP isoprene experiments.

From this study it is hard to confirm the role of *Lysobacter* in isoprene degradation. *Lysobacter* is significantly increased in many of the serial enrichment microcosms, has a variable SIP profile, with significance in one heavy fraction indicating any interaction being indirect, and a low concentration experiment result significant enrichment of birch soil, but at such a low level it could be just barely present in the test replicates only, alongside increasing means in the willow set. Unfortunately apart from sporadic enrichment resulting somehow from high isoprene, there is nothing concrete that can be derived from these data. Summporting this; *Lysobacter* abundance changes with

differing concentrations of isoprene were not significant ($F_{8,62}$ = 1.39 P=0.22) in the enrichment experiments, and *Lysobacter* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 0.11 P=1) in the low concentration isoprene experiments. *Lysobacter* abundances at different densities were not significant ($F_{3,65}$ = 1.68 P=0.18), and with carbon types were not significant ($F_{3,65}$ = 0.85 P=0.47) in the SIP isoprene experiments.

Variovorax species seem to be enriched mainly in the 7.2 $\times 10^5$ ppb third level enrichment of the serial enrichments (with seemingly three OTUs involved), with little presence at 7.2 $\times 10^6$ ppb, however they are significantly decreased in the heavy fractions of DNA SIP, and show no significant changes in the low concentration experiments. Taken together this information indicates the likelihood that whatever effect causes the enrichment of *Variovorax* with isoprene, is firstly dependant on isoprene concentration with inhibition at high concentrations, and secondly is not *Variovorax* metabolising isoprene. *Variovorax* however, was isolated on minimal media with isoprene as a sole added carbon source (Chapter 3), so any factor causing this unusual enrichment profile may not be biological. This is in stark contrast to the El Khawand (2014) study, which found *Variocorax* only in the heavy fractions of isoprene SIP, although this lacked replication, statistics, and could have been a different species. Without replication, and with *Variovorax* having GC content of 66.5%, in a bacterial GC content range of 16 to 75%, the DNA is comparatively dense, so it's presence in the heavy fraction could have been entirely stochastic (Han et al., 2013; Lightfield et al., 2011). The varying DNA density between bacteria, and therefore fraction location is an inherent problem in using just two

fractions for DNA SIP; and is why testing for a heavy shift throughout a range of fraction densities for each OTU is superior. *Variovorax* abundance changes with differing concentrations of isoprene were not significant ($F_{8,62}$ = 1.17 P=0.33) in the enrichment experiments, and *Variovorax* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 1.08 P=0.43) in the low concentration isoprene experiments. *Variovorax* abundances at different densities were not significant ($F_{3,65}$ = 0.86 P=0.47), and with carbon types were not significant ($F_{3,65}$ = 0.82 P=0.49) in the SIP isoprene experiments.

Pedobacter has been isolated on minimal media with isoprene as the sole added carbon source, is significantly enriched in 9 of the sequential enrichments across 5 of the soil types, with four OTUs, and increases, although insignificantly, in the low concentration experiments, and although nothing can be garnered from the DNA SIP, which is has no clear pattern or any significance, for these reasons it is possible that it is involved in isoprene degradation, certainly at high concentrations, possibly at low, but it is cannot be determined from these data if it is direct or indirect. *Pedobacter* abundance changes with differing concentrations of isoprene was significant ($F_{8,62}$ = 3.55 P= < 0.001) in the enrichment experiments, and *Pedobacter* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 0.01 P=1) in the low concentration isoprene experiments. *Pedobacter* abundances at different densities were not significant ($F_{3,65}$ = 0.29 P=0.84), and with carbon types were not significant ($F_{3,65}$ = 0.59 P=0.63) in the SIP isoprene experiments.

Ramlibacter species were enriched in the sequential enrichments, particularly in the 7.2 $\times 10^5$ ppb third level enrichments, suggesting a potential interaction with isoprene, however the lack of presence in the DNA SIP experiments and the lack of change in the low concentration experiments suggest that any effect, even if direct, is unlikely to be biologically relevant. *Ramlibacter* abundance changes with differing concentrations of isoprene was significant ($F_{8,62}$ = 5.46 P= < 0.001) in the enrichment experiments, and *Ramlibacter* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 0.73 P=0.67) in the low concentration isoprene experiments. *Ramlibacter* abundances at different densities were not significant ($F_{3,65}$ = 1.09 P=0.36), and with carbon types were not significant ($F_{3,65}$ = 0.57 P=0.64) in the SIP isoprene experiments.

Mucilaginibacter, despite meeting the inclusion threshold of two statistical significances on enrichment, is almost certainly not involved in isoprene degradation, with as much significance of suppression through enrichment, as well as a significant decrease at low concentration (not shown). It is likely that the enrichment is due to highly stochastic processes, even if only occasionally replicable and sometimes affected by isoprene. *Mucilaginibacter* abundance changes with differing concentrations of isoprene were not significant ($F_{8,62}$ = 0.88 P=0.54) in the enrichment experiments, and *Mucilaginibacter* abundances at different densities were not significant ($F_{3,65}$ = 2.45 P=0.07), and with carbon

types were not significant ($F_{3,65}$ = 1.66 P=0.19) in the SIP isoprene experiments.

Aeromonas species do not have strong evidence of isoprene degradation, with some significance in early enrichments, only one of which has a triplicate mean above 4%, and no presence in DNA SIP or the low concentration experiments. However, adding these modest increases to the presence and (also modest, with no significance stated) increase shown in the heavy fraction of isoprene DNA SIP in marine systems (Johnston., 2014), it is circumstantially plausible, that some *Aeromonas* species may be involved in isoprene degradation. *Aeromonas* abundance changes with differing concentrations of isoprene were not significant ($F_{8,62}$ = 0.39 P=0.93) in the enrichment experiments, and *Aeromonas* abundances at different densities were not significant ($F_{3,65}$ = 1.87 P=0.14), and with carbon types were not significant ($F_{3,65}$ = 2.31 P=0.09) in the SIP isoprene experiments.

Shinella species are unlikely to be involved in isoprene degradation in these soils, at least directly, although enriched in some serial enrichments, DNA SIP evidences their lack of a role in direct isoprene metabolism, however a "*Shinella*-like" organism was described as being involved in isoprene degradation in marine systems (Acuña Alvarez et al., 2009), although this could be a different species, it could have also been enriched by whatever indirect mechanism caused the enrichment of *Shinella* in these experiments. *Shinella* abundance changes with differing concentrations of isoprene was significant

($F_{8,62}$ = 4.23 P= < 0.001) in the enrichment experiments, and *Shinella* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 0.01 P=1) in the low concentration isoprene experiments. *Shinella* abundances at different densities were not significant ($F_{3,65}$ = 1.13 P=0.35), and with carbon types were not significant ($F_{3,65}$ = 0.31 P=0.82) in the SIP isoprene experiments.

Bosea species, despite being significantly enriched in some serial enrichments, and having a species isolated as part of this work (Chapter 3), are also suppressed in others, not present in DNA SIP and have no enrichment in low concentration experiments, indicating they are unlikely to be involved in isoprene degradation. *Bosea* abundance changes with differing concentrations of isoprene was significant ($F_{8,62}$ = 3.19 P=0.01) in the enrichment experiments, and *Bosea* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 0.01 P=1) in the low concentration isoprene experiments. *Bosea* abundances at different densities were significant ($F_{3,65}$ = 5.08 P= < 0.001), and with carbon types were significant ($F_{3,65}$ = 0.32 P=0.82) in the SIP isoprene experiments.

Acetonema seem to be enriched in many of the 0.1 ml 30°C isoprene headspace serial enrichments (7.2 ×10^5 ppb), however in none of the 7.2 ×10^6 ppb, suggesting inhibition by, or out competition in high isoprene levels. *Acetonema* species are not present in DNA SIP experiments, and are not enriched in low concentration experiments. *Acetonema* abundance changes

with differing concentrations of isoprene were not significant ($F_{8,62}$ = 1.41 P=0.21) in the enrichment experiments, and *Acetonema* abundance changes with differing concentrations of isoprene was significant ($F_{8,15}$ = 3.39 P=0.02) in the low concentration isoprene experiments. *Acetonema* abundances at different densities were not significant ($F_{3,65}$ = 2.14 P=0.1), and with carbon types were not significant ($F_{3,65}$ = 0.5 P=0.69) in the SIP isoprene experiments.

Polaromonas, *Caenimonas* and *Acinetobacter* species are unlikely to be involved in isoprene degradation, despite some significant increases in sequential isoprene enrichments, these are generally sporadic, and the relative abundance is generally quite low (<0.15%), with no additional evidence of degradation from low concentration experiments or from DNA SIP, it is hard to tell from these experiments.

Methylobacterium showed no significant enrichment in the sequential enrichments, however the DNA SIP shows a heavy shift (although not significant), a *Methylobacterium* representative has been cultured on minimal media with isoprene as the only provided carbon source, and *Methylobacterium* shown an enrichment in the low concentration experiments (although not significant), from near zero to over 2% of sequences. This demonstrates that this *Methylobacterium* includes isoprene degraders. Although considering the lack of increase at high concentrations, it is likely that *Methylobacterium* is not important in degrading, or is inhibited by high concentrations of isoprene, it is capable of isoprene degradation in concentrations similar to environmental conditions, and is therefore likely important *in situ*. *Methylobacterium*

abundance changes with differing concentrations of isoprene were not significant ($F_{8,62}$ = 0.32 P=0.96) in the enrichment experiments, and *Methylobacterium* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 1.07 P=0.43) in the low concentration isoprene experiments. *Methylobacterium* abundances at different densities were significant ($F_{3,65}$ = 2.72 P=0.05), and with carbon types were significant ($F_{3,65}$ = 0.77 P=0.52) in the SIP isoprene experiments.

Nakamurella, similarly did not show any significance in sequential enrichments, and had a significant enrichment in the low concentration experiments; however DNA SIP experiments show only a relative abundance shift towards the light fractions (with no significance), so it is unlikely to be involved in direct degradation, if it is involved in degradation at all. *Nakamurella* abundance changes with differing concentrations of isoprene were not significant ($F_{8,62}$ = 0.6 P=0.78) in the enrichment experiments, and *Nakamurella* abundance changes with differing concentrations of isoprene were not significant ($F_{8,15}$ = 0.63 P=0.74) in the low concentration isoprene experiments. *Nakamurella* abundances at different densities were significant ($F_{3,65}$ = 3.8 P=0.01), and with carbon types were significant ($F_{3,65}$ = 0.3 P=0.83) in the SIP isoprene experiments.

2.6 Conclusion

Although the in situ data from this project only represents a small area of Essex, and a few days across the year, two firm conclusions can be drawn. Firstly that if there is a seasonal effect, it is not a significant driver of rate, and secondly, that the current figure we have for global terrestrial isoprene degradation is incorrect, and the rates reported in Gray et al. (2015) are more representative of reality.

The lack of seasonal trend (and the seasonal effects of leaf cover) also supports the lack of difference between the isoprene degradative capacity in soils away from or from near isoprene producing trees. In addition to the differential ability of some isoprene degraders to take advantage of high levels of isoprene, and the very rapid degradation of isoprene by soils communities, although it is possible that isoprene concentrations were insufficiently localised in an open park, it is likely that something other than nearby isoprene producing trees is supplying isoprene to the soil, possibly an internal soil isoprene cycle.

This work has confirmed that *Rhodococcus*, particularly R. *globerulus*, *wratislaviensis*, *erythropolis*, *opacus* are key isoprene degraders at very high levels of isoprene and are clearly primed to take advantage of high isoprene concentrations, however has cast a shadow of doubt on their involvement at realistic concentrations of isoprene. However, it has also indicated some groups, like *Methylobacterium* and possibly *Fluviicola* which may be involved in isoprene degradation in situ, and raised the possibility of a multi-tiered isoprene degradation situation, with different bacteria ranging widely in affinity for isoprene. It has also identified some groups which are enriched due to isoprene, but are unlikely to be direct degraders, with TM7, *Aeromicrobium*,

Nocardioides and *Acidovorax* being the most notable, and identified the OTUs affected, and their most similar species.

Overall the approaches used to investigate the communities involved in isoprene degradation made up a quite robust set of experiments, with the sensitivity of sequential, media rich high isoprene enrichments, the direct evidence of SIP and the media free environmental isoprene levels in the low concentration experiments – across two different MiSeq and bioinformatics runs, across two years. Where evidence from these multiple sources suggests bacterial involvement it is strong. That said, in retrospect, the serial enrichments would have been better performed at lower concentrations, to reflect in situ conditions, the earlier termination of the Stable Isotope Probing experiment may have been more sensitive, or a lower concentration (if possible with the DNA requirements) could have been advised, and the low concentration enrichments could have either been continued for longer, or could have done with a greater number of replicates in order to the increase sensitivity of statistical tests to small changes.

Certainly from this work comes the conclusion that any further work on isoprene degradation in the environment should focus on low, environmentally relevant, concentrations of isoprene.

Many thanks to Toungphorn Uttarotai (visiting student) for her assistance with the DNA SIP lab work.

NB: Samples were also taken, and serially enriched from the lakewater in Wivenhoe park, and some of the serial enrichments were 454 pyrosequenced, and all sequential enrichments were amplified and analysed by DGGE, and band sequencing (with similar, but lower resolution conclusions), however as these are tangential/ superseded to/in this project, they have not been included.

Chapter 3. Isolation of bacteria associated with isoprene degradation and investigation of the isoprene degradation pathway.

3.1 Abstract

Isoprene is a climate altering, health affecting, volatile hydrocarbon. Bacteria that are involved in the biodegradation of isoprene have been identified previously, and the genetic regions involved in isoprene degradation have been characterised in a few species, with early forays into investigating the wider diversity of isoprene degrading genes in isolates and in the environment having been performed using primers based on this information. This chapter describes the isolation and characterisation of numerous bacterial strains associated with isoprene degradation, and the investigation of their phylogeny and putative isoprene monooxygenase genes. Comparing the *isoA* and *16S rRNA* gene sequences of bacteria revealed that there was no correlation between *isoA* sequences and *16S rRNA* sequences, and three distinct clusters of *isoA*-like genes were identified, only one of which was similar to previously studied sequences. Selected, strains, which were confirmed to degrade isoprene and grow on it as sole carbon source, were subjected to carbon source testing, and/or genomic sequencing, annotation and investigation, in order to increase the breadth and depth of our knowledge of the metabolic activity of isoprene degraders and the genetics associated with isoprene degradation, demonstrating that the isoprene degrading operon was present in different genomic locations, with different gene orientations, in closely related bacteria. The metagenome of a TM7 rich isoprene degrading community was also sequenced to elucidate the mechanisms behind its isoprene driven enrichment, revealing a possible metabolic interaction between TM7 and *Aeromicrobium* that merits further investigation.

3.2 Introduction

One-third (600 TgC) of the world's annual hydrocarbon production is the highly reactive 2-methyl-1,3-butadiene, commonly known as isoprene. Isoprene is mostly produced by trees and shrubs as a method of reducing damage by transient heat (Sharkey et al., 2008), and is mostly attenuated through atmospheric reactions – often producing health or climate affecting compounds (Fan & Zhang, 2004). For example, reactions with free radicals reduce their atmospheric concentration resulting in less capacity for free radicals to react with other hydrocarbons, which increases the retention time of the potent greenhouse gas, methane. Other reactions with nitrous oxides produce tropospheric ozone, which can cause respiratory problems (Ashworth et al., 2013; Jenkin & Clemitshaw, 2002). Isoprene is highly reactive due to its carbon-carbon double bonds, leading to an atmospheric lifetime of less than one hour, and a resulting atmospheric concentration between the low parts per trillion where production is low (for example, above marine environments), up to over 10 ppb in tropical regions (Seinfeld & Pandis (2012); Baldocchi et al., 1995; Kesselmeier et al., 2000; Lewis et al., 1997).

However, evidence is emerging that a minority of isoprene, possibly up to 5%, is degraded by soil dwelling microbes (Cleveland & Yavitt, 1997). It has already been established that representatives of genera, including *Rhodococcus*, *Mycobacterium* and *Gordonia,* are capable of degrading isoprene (Johnston., 2014; El Khawand, 2016; Acuña Alvarez et al., 2009; van Hylckama Vlieg et al., 2000); with evidence that some fungi are involved as well (Gray et al., 2015). Work by the group of Janssen (e.g. van Hylckama Vlieg et al., 2000) and the

group of Murrell (e.g. El Khawand, 2014); Johnston, 2014; Crombie et al., 2015) have given insight to the mechanisms of isoprene degradation in *Rhodococcus*, *Mycobacterium*, and *Gordonia* species, and resulted in primers and protocols designed for the investigation of the isoprene degradation pathways.

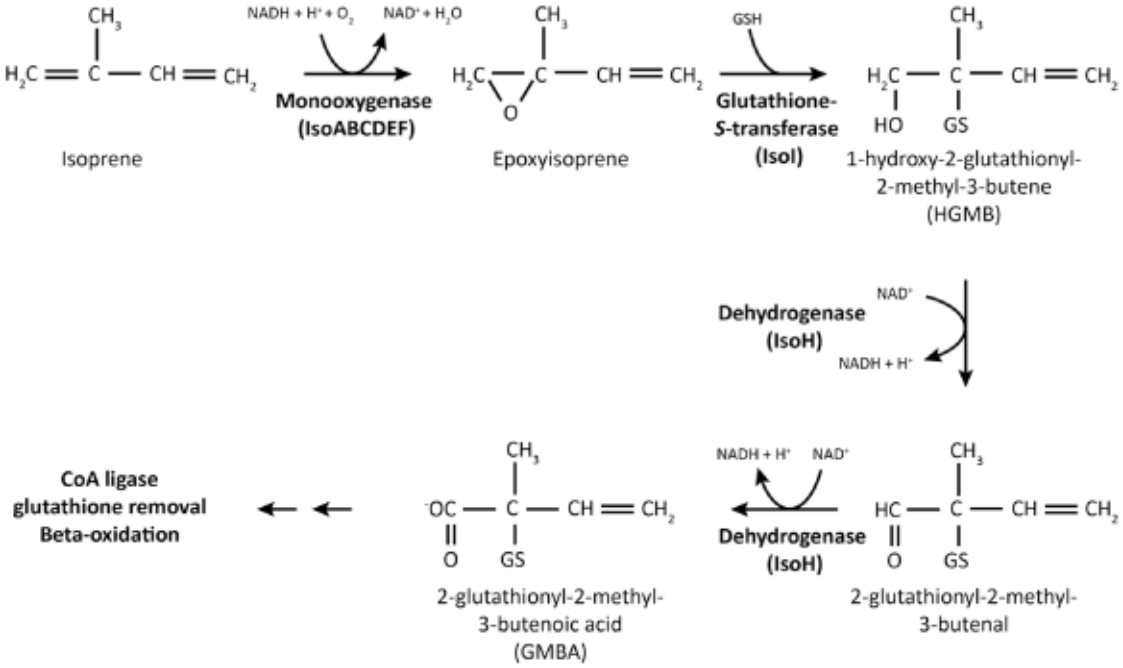


Figure 3.1. Isoprene degradation pathway of *Rhodococcus* AD45 from van Hylckama Vlieg et al. (2000).

Isoprene degradation is thought to take place through enzymes *isoA* to J, encoded on a megaplasmid by genes *isoA* to *J*. Isoprene degradation begins with a monooxygenases, *isoA*BCDEF, where isoprene is converted into epoxyisoprene. A glutathione-*S*-transferase, IsoI, converts this into 1-hydroxy-2-glutathionyl-2-methly-3-butene (HGMB). A dehydrogenase, IsoH, converts HGMB into 2-glutationyl-2-methyl-3-butenal and then through to 2-glutathionyl-2-methyl-2-butenoic acid (GMBA) ((van Hylckama Vlieg et al., 2000) (Figure 3.1).

Using '*isoA* primers' targeting the *isoA* gene El Khawand (2016) and Johnston (2014) were able to characterise a number of *isoA* sequences from environmental samples, giving an initial insight into the diversity of *isoA* sequences in the environment.

The aim of this chapter was to isolate a wide range of isoprene degrading bacteria in order to: (i) understand the phylogenetic and physiological breadth of microbes involved in the process, (ii) compare the sequences of the key phylogentic (16S rRNA) and key functional (*isoA*) genes to gain insights into the evolution of isoprene degradation, and (iii) Analyse the genomes of selected isoprene degrading strains to further understand the diversity of the isoprene degradation pathways. Having ascertained that TM7 was highly enriched isoprene fed soils, another aim was to isolate a representative of this phylum and failing that to investigate it by metagenomic analysis of an enrichment

3.3 Materials and Methods

3.3.1 Enrichment and isolation of isoprene degrading microorganisms

Six soil, five leaf and one lake location were sampled from Wivenhoe Park (UK) in triplicate in early October 2012; each in triplicate. The leaves used in this study were from the following trees: Poplar (*Populus canadensis*), Willow (*Salix babylonica*), Oak (*Quercus robur*) Birch (*Betula pendula*) and Ash (*Fraxinus excelsior*), with the triplicates spread equally around the circumference of the same tree, and taken at head height (or nearest available height). Soil samples were taken within one meter of the tree trunk of each of these trees, and between one and two meters from the other replicates (see Section 2.3.1 for

details). Lake water was sampled from a lake near the Ash location (51°52'42.2"N, 0°56'53.8"E). Leaf samples consisted of entire leaves, which were encompassed by sterile whirlpack bags and cut off at the stem. Soil samples consisted of the un-sieved top two centimetres of vegetation free soil (See Section 2.3.1 for details), and lake water samples consisted of 50 ml from the top 10 cm of the lake, 0.3 m from the edge, sampled with Falcon tubes. Microbes living on the surface of the leaves were detached using a modified version of a protocol provided by Aslam (2012); a mixture of large and small whole leaves were suspended in 0.02 M Tris, 0.01 M EDTA, 0.024% v/v Triton X solution (final concentrations from a 50×, pH 7.5 stock) and shaken for 30 min, and then sonicated for 10 min, and filtered through sterile glass wool. One ml of lake water, 1 ml of cell suspension washed from leaves or 1 g of soil (as applicable) of each sample was incubated in 9 ml minimal medium in a sterile 125 cm^3 serum bottle with a PTFE covered silicon septum, and isoprene added to final concentrations of 7.2×10^5 and 7.2×10^6 ppb, (as in Section 2.3.2.1). The isoprene concentration in the microcosms was measured by GC-FID, and after observing over 80% decreases in the isoprene concentration, 1 ml of the sample was set up in a new enrichment, with this sequential enrichment two times, as in (as in Section 2.3.2.1).

The third enrichment for each triplicate was diluted 10^1, 10^2, 10^3 and 10^4 fold in the same minimal media, and 100 µl of each dilution was streaked onto minimal media plates containing 15 g l^{-1} agar, and incubated in glass desiccation jars containing approximately 0.5 ml of 99% isoprene per 1000 cm^3 of air for 7 days. Unique looking colonies were picked, and serially re-plated and incubated for

isolation. Numerous strains were also isolated through a number of different methods, outlined in Appendix 3.AX.3. From these, 141 strains were chosen for partial sequencing of their 16S rRNA gene, and 95 for partial sequencing of their *isoA* gene, based on different colony morphologies and an ability to grow on isoprene as the sole added carbon source. A list of strains and descriptions is available in Appendix 3.AX.1. A small quantity of cells from a single colony was lifted using a 10 µl pipette tip, and suspended in 20 µl sterile DNAase free water. The DNA was extracted by freezing at -20°C, followed by 5 minutes of heating to 95°C in a thermocycler.

Following DNA extraction, the *16S rRNA* and the *isoA* genes were amplified by PCR: 25 µl RedTaq mastermix, 2 µl DNA solution, 1 µl forward primer, 1 µl reverse primer and 21 µl water were added to each well of a 96 well plate, sealed, and spun briefly. For *16S rRNA* amplification the primers used were 27F (3-AGAGTTTGATCMTGGCTCAG-5) and 1492R (3-CGGTTACCTTGTTACGACTT-5) (Weisburg et al., 1991), and PCR was performed with a 5 minute hold at 94°C, followed by 35 cycles of 94°C for 30 s, 56°C for 30 s and 72°C for 90 s, with a 10 minute final 72°C extension. *isoA* gene amplification was performed using the *isoA* primers and protocol developed by UEA (Crombie et al., 2015), *isoA*-F (3-TGCATGGTCGARCAYATG-5) and *isoA*-R (3-GRTCYTGYTCGAAGCACCACTT-5), with a hold of 94°C for 4 min, followed by 19 cycles of; 94°C for 30 s, 72°C (decreasing by 1°C per cycle) for 45 s and 72°C for 60 s, and 25 cycles of 94°C for 30 s, 54°C for 45 s and 72°C for 60 s, with a final 72°C hold of 5 minutes. PCR products were purified by addition of 50 µl Ampure XP magnetic beads, incubating for 5 minutes, on a magnetic

stand (Agencourt SPRiPlate), removal of the supernatant, washing twice with 190 µl of 80% v/v ethanol, air drying for 10 minutes, removal from the magnetic plate and suspension in EB (Elution Buffer: 10 mM Tris-Cl, pH 8.5) for two minutes, followed by incubation on the magnetic plate for 30 seconds and recovery of the EB containing purified DNA. Sanger sequencing was performed by GATC Life Sciences using the lightrun plate option and the forward primers.

3.3.2 Analysis of the 16S rRNA and *isoA* gene sequences of isoprene degrading strains

Sanger sequences were quality filtered and converted to FASTA format using Chromatogram Explorer v3.3 from Heracle Biosoft, with an end trimming quality value limit of 18, and a window length of 10 for 75% good bases.

FASTA sequences were aligned using the MUSCLE algorithm (default parameters) in MEGA (6.06), and pairwise distance matrices were calculated using the distance function in MEGA with the Jukes and Cantor model (default parameters, nucleotide sequences).

Pairwise distances were exported from MEGA into MS EXCEL (2013), and aligned, renamed, and sorted, to give an *isoA* and a *16S rRNA* distance (base substitutions per site) for each pair of strains. Analysis was carried out by graphical representation in MS EXCEL (2013), and tests of normality (Shapiro-wilk, linear regression, residual analysis), linear regression, and Spearmans rank correlation in *R.*

The closest relatives to strains were identified by BLAST-n and type strains' sequences were obtained from LPSN, and the type strains *16S rRNA* gene sequences were aligned using the same methodology, and used to generate a

maximum likelihood phylogenetic tree in MEGA 6. Once identified by the equivalent *16S rRNA* gene sequences the *isoA* sequences were aligned, and used to create a maximum likelihood phylogenetic tree.

3.3.3 Carbon source testing

Selected strains were incubated in Biolog plates PM1 and PM2 for 48 hours, according to the Biolog gram positive plate instructions, using the high growth dye for all strains apart from strain GM3 (slow growth dye). Growth was measured using a plate reader (labtech FLUOstar Omega) to measure the Tetrazolium based respiration indicative dye as a proxy for respiration at 590 nm. Volatile hydrocarbon testing plates were set up in (empty) 96 well plates with the same biolog mix, the plates were sealed and incubated in desiccation jars with 0.2 ml isoprene per 1000 cm^3 air, toluene, benzene, hexane, ethane, DMSO, or flushing with methane, propane or propane, or nothing, and respiration was measured in the same way (with Tetrazolium dye, and read on a plate reader) as the Biolog plates.

3.3.4 Genome sequencing

DNA from representative strains, which were confirmed to degrade isoprene, was extracted using the GenElute bacterial Genomic DNA Kit (Sigma), with no protocol modifications. DNA integrity was checked on a 0.5% agarose gel, and quantity was measured by spectrophotometry (Nanodrop 3300 fluorospectrophotometer with PicoGreen dye). The genomic DNA was sequenced using PacBio technology at NBAF-L, and was kindly assembled into contigs by Xuan Liu at NBAF-L using PacBio's SMRT-Analysis software 2.3.0

with a minimum subread length of 50, minimum polymerase quality of 75, and a minimum polymerase read length of 50. Bespoke VBA code was used to convert the assembled contigs into multiFASTA format, and the genomes were uploaded to RAST (Aziz et al., 2008; Brettin et al., 2015; Overbeek et al., 2014) for automatic annotation (default parameters) and manual investigation.

Differential predicted metagenomic interaction analysis was performed on enrichment b22ck consisting of around 35% 16S rRNA gene sequences derived from TM7. The community was sequenced by PacBio, and assembled into contigs as above. Each contig was identified to the genus level by BLAST-n (Altschul et al., 1990). A bespoke algorithm was created to produce two separate multiFASTA files, one with and one without TM7. Both sets of sequences were uploaded to MGRAST (Meyer et al., 2008) for function prediction, and their predicted functions were overlaid on a KEGG pathway map (Kanehisa & Goto, 2000). Where TM7 biochemical pathways were present, but were absent in the rest of the community, it was assumed that this was likely a unique function of TM7 in the community. The proteins predicted to be involved in supplying or utilising these metabolites were deemed likely to belong to groups metabolically interacting with TM7, and the protein sequence was identified from the KEGG database (from *E. coli*). This protein sequence was searched against a local BLAST database constructed with the metagenomic community using tBLASTn (through the prfectBLAST frontend (Santiago-Sotelo & Ramirez-Prado, 2012)), and the matching metagenome contigs were identified. These contigs were related back to the original contig BLAST-n result to indicate which members of the microbial community theoretically interact with

TM7 through their metabolic activity. The functions predicted by MGRAST for the TM7 and non-TM7 metagenomic data were also investigated for any differences in antibiotic resistance, or virulence factors; however there =were no obvious differences.

3.3.5 Attempts at isolating TM7

Many additional approaches for isolating bacteria that have proven recalcitrant to cultivation, particularly TM7, were deployed for enrichments derived from one soil sample taken from beneath a willow tree. Techniques included using different surface types (e.g. agar replacements), carbon sources and antimicrobials, and providing direct or indirect access to metabolites from other bacteria. More details are available in the Appendix (3.AX.3).

3.4 Results

3.4.1 Numerous types of bacteria are able to grow on isoprene as the sole carbon source.

In order to investigate the culturable isoprene-degrading microbiota, isoprene enriched communities were plated onto minimal media with isoprene as the sole carbon source. Colonies with distinct morphologies were picked and subcultured until the same morphology only was visible for over six generations. The partial *16S rRNA* gene was sequenced for these isolates, and the closest species were identified by BLAST-n.

Table 3.1 Table categorising the number of isolates belonging to (or most close to) genera based on BLAST-n analysis of 16S rRNA gene sequences for the isolates used in the phylogenetic analyses.

Genus	Quantity	Genus	Quantity
Aeromicrobium	**1**	***Nocardia***	**1**
Arthrobacter	**4**	***Nocardioides***	**1**
Bosea	**1**	***Pedobacter***	**3**
Dietzia	**1**	***Pseudomonas***	**2**
Ensifer	**1**	***Rhodococcus***	**105**
Mezorhizobium	**1**	***Stenotrophomonas***	**3**
Methylobacterium	**2**	***Variovorax***	**2**
Nitrobacter	**1**		

Refer to Appendix 3.AX for detailed information on the isolates generated in this project

The vast majority of isolates belong to the genus *Rhodococcus*, particularly *R. erytheropolis*, *R. globerulus* and *R. opacus* (Table 3.1).

In order to accurately visualise the phylogenetic relationship of these isolates, type strains for each of the top three highest identity BLAST-n results stating the species level were obtained from the LPSN, and the type strains 16S rRNA gene sequence was used to make a maximum likelihood phylogenetic tree.

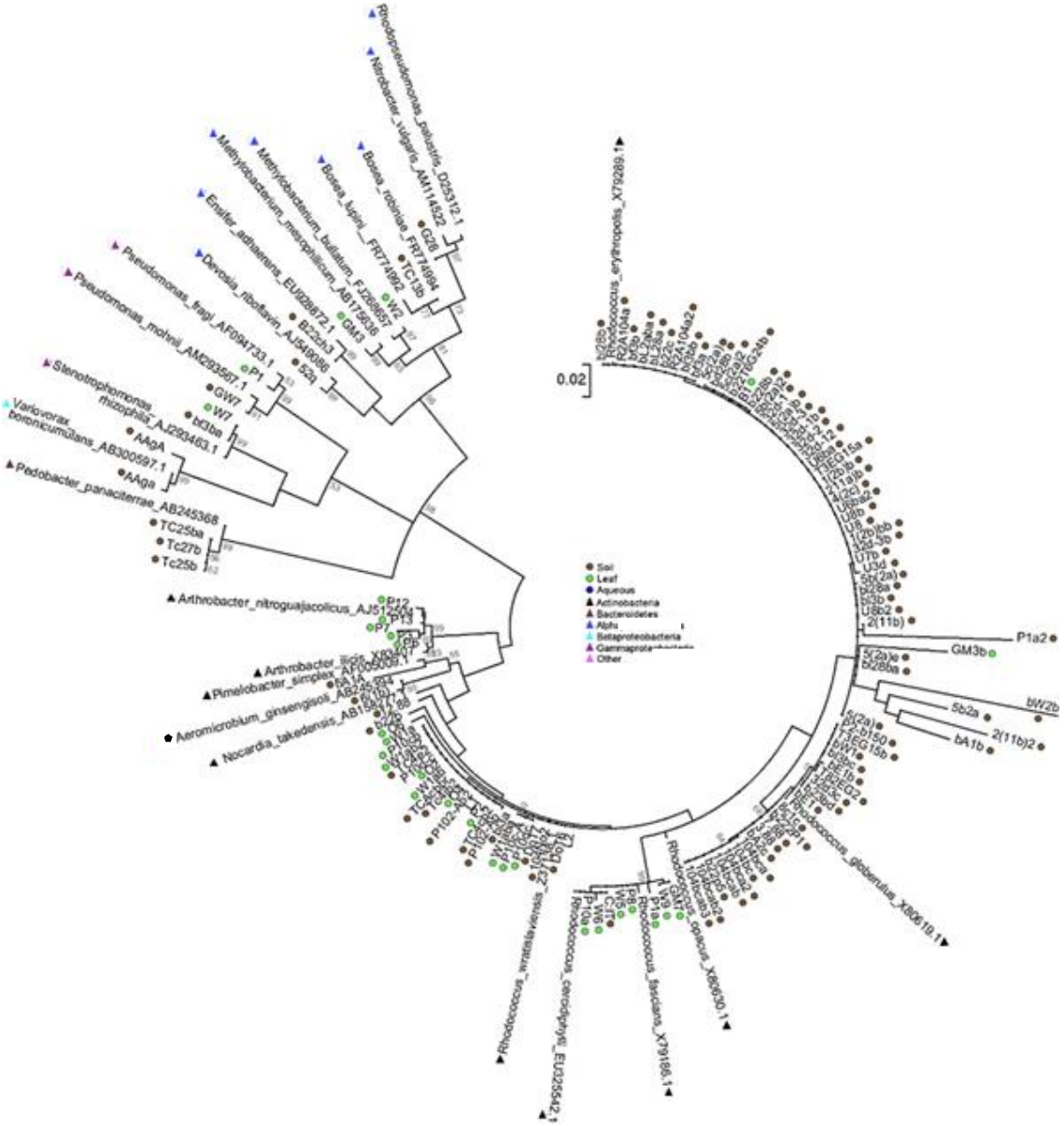


Figure 3.2 16S rRNA gene sequence maximum composite likelihood tree following MUSCLE alignment (50 iterations) and using the Tamura-Nei (1993) model and the highest log likelihood. Initial tree generation was through the Jukes and Cantor Neighbour Joining and BioNJ models (MEGA 6). (950 bp sequences). Bootstrap resampling (1000 replications) was used to test the validity of the branching order and values over 50% are reported (grey). Type strains were identified through searching LPSN for the names of nearest BLAST results. External circles mark the origin of the strain , green = leaves, brown = soil. External triangles on type strains indicate the phylum or sub-phylum: Black: Actinobacteria, Brown: Bacteroidetes, Dark blue: Alphaproteobacteria, Light blue: Betaproteobacteria, Purple: Gammaproteobacteria. The scale shows nucleotide substitutions per site for radial distance.

The phylogenetic tree broadly agrees with the blast-n results, and displayed a high diversity of organisms culturable with isoprene as the sole carbon source. Strain G28, initially indicated to be *Nitrobacter* is equally close to *Rhodopseudomonas*, and strain 5A1A is not particularly close to the *Pimelobacter* which is the nearest type strain, potentially representing a new species. Likewise, on the far right, there are a number of strains which are related to, but divergent from *Rhodococcus*. Clear from Figure 3.2 is that there

3.4.2 There is no correlation between *isoA* gene sequence and *16S rRNA* sequence

As little is known about the diversity of *isoA* gene sequences, nor how they are related to phylogeny, 93 isolates (selected due to having both *isoA* and *16S rRNA* sequences meeting the quality cut offs in section 3.3.2) were investigated for *isoA* gene sequences, as well as their *16S rRNA* gene sequences, which are predictive of phylogeny. *isoA* sequence diversity was investigated by creating a phylogenetic tree of these sequences, along with the *isoA* sequences from El Khawand (2016) (Figure 3.3), which shows that most of the El Khawand (2016) sequences clustered together, along with a minority of the sequences from this project, indicating that there is a greater diversity of *isoA* products than previously supposed.

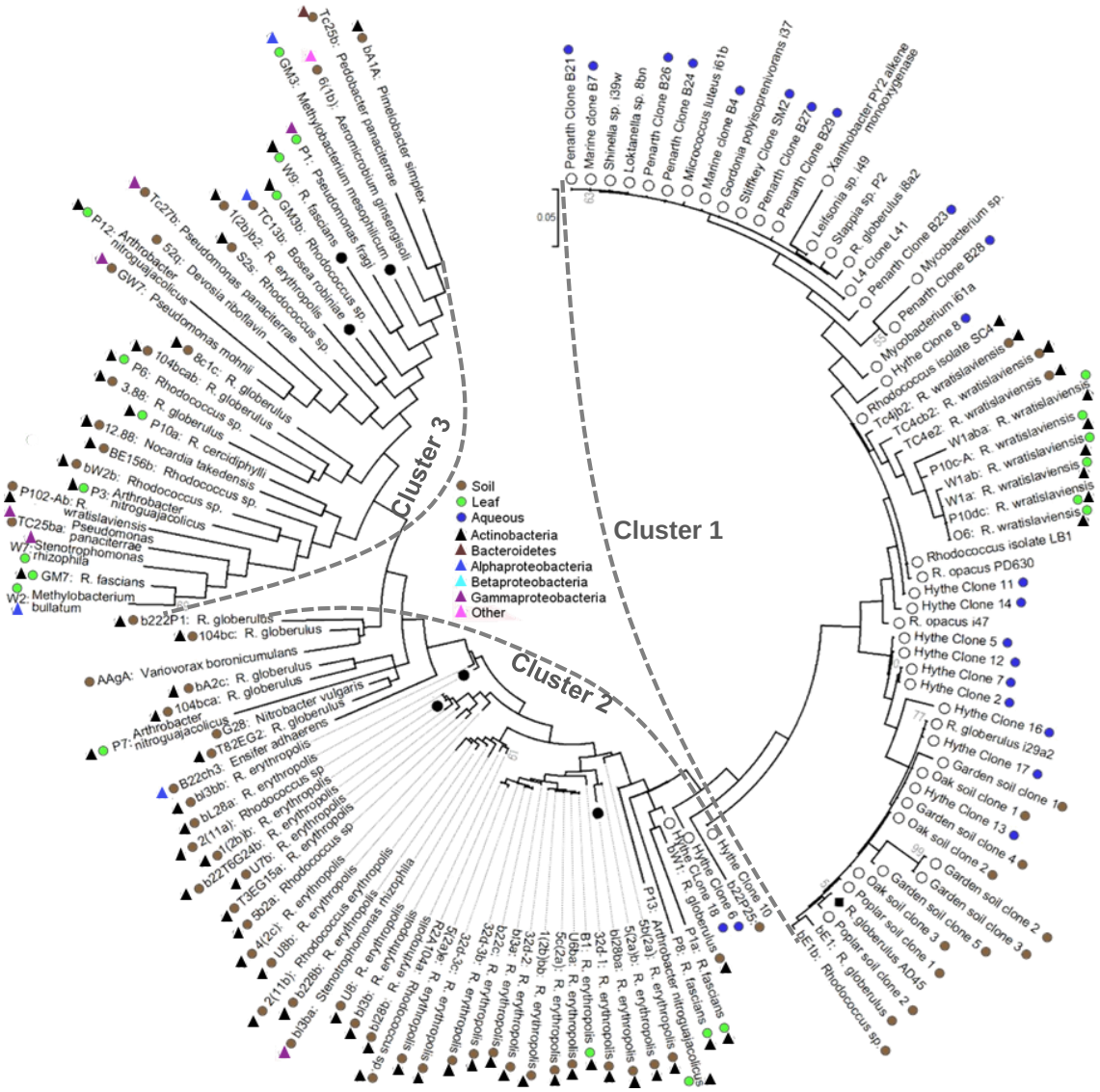


Figure 3.3 Phylogenetic tree based on *isoA* nucleotide sequences (maximum composite likelihood tree using the Tamura-Nei (1993) model and the highest log likelihood). Initial tree generation was through the Jukes and Cantor Neighbour Joining and BioNJ models (MEGA 6). Strains are labelled with their nearest type strain based on 16S rRNA gene sequence analysis. At the branch termini: sequences from strains and environmental clone libraries from El Khawand (2016) are marked by white circles, strains used for further investigation are marked by internal black circles, and *Rhodococcus* AD45 is marked by black diamond. Circles at the end of the strain name indicate the origin of the strain: green = leaves, brown = soil, blue = (aqueous/semi) estuary/sediment. External triangles mark phylum or sub-phylum for strains from this thesis: Black = Actinobaceria, Brown = Bacteroidetes, Dark blue = Alphaproteobacteria, Light blue = Betaproteobacteria, Purple = Gammaproteobacteria. "R." denotes *Rhodococcus.* 830 bp sequences was used. The main clusters are marked by dotted lines. The scale bar shows the number of nucleotide substitutions per site for radial distance. Bootstrap values (x1000 replicates) over 50 are displayed as gray numbers near branch nodes.

The diversity of *isoA* sequences from this thesis was much greater than the diversity previously reported (e.g. Cluster 2 with the strains from El Khawand et al. 2016), with only a little overlap with the *isoA* diversity previously reported. There was generally little correspondence between phylogeny based on *16S rRNA* gene sequences and *isoA* gene sequences, for example, there are *Rhodococcus* representatives present in each cluster, and *Arthrobacter* and *Stenotrophomonas* both have strains in Cluster 2 and 3. Interestingly, *Rhodococcus globerulus* AD45 grouped with *R. globerulus* i29a2 from El Khawand (2016), and *R. globerulus* 104a from this project. Cluster 2 sequences had around 70% similarity to the *Rhodococcus* AD45 *isoA* sequence on BLASTn analysis (with no other results), however Cluster 3 had no results. As some *isoA* sequences seem to group according to their phylogeny (assuming here that 16S rRNA genes represent the true phylogenetic position), the hypothesis of a relationship between *isoA* gene sequence and phylogeny based on 16S rRNA gene sequences was tested. In order to test this, similarity matrices were calculated to generate distances for each of 92 isolates (which have sufficient quality sequence for *16S rRNA* and *isoA* products) against each of the other isolates for both the *16S rRNA* gene, a common proxy for phylogeny, and the *isoA* gene products. These *16S rRNA* distance pairs were matched with the *isoA* distance pairs for the corresponding pairs, and the *16S rRNA* distances were plotted against the *isoA* distances, allowing investigation of the relationship.

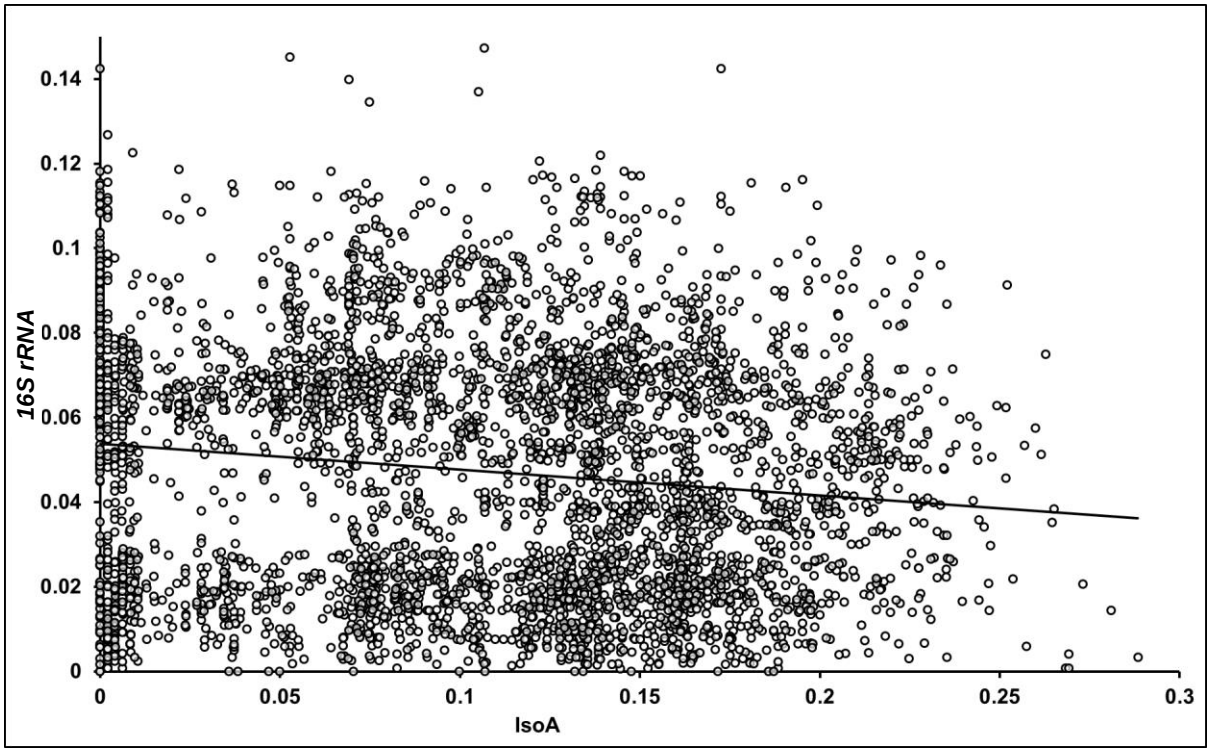


Figure 3.4 Pairwise distances for *isoA* gene sequences against 16S rRNA gene sequences sequence for 92 strains (4005 pairs) using MUSCLE alignment and the Jukes and Cantor model. The line of best fit is shown and $R^2 = 0.0184$.

The distribution of pairwise points (Fig. 3.4) supports the observations from the *isoA* gene tree (Fig. 3.3), that there is no correlation ($R^2 = 0.0184$) between 16S rRNA and *isoA* gene sequences of the isolates.

3.4.3 Isoprene degrading bacteria have a wide range of other metabolic capabilities

In order to understand the metabolic context of isoprene degradation, six strains were chosen for carbon source testing. Isolates were screened for growth on a variety of carbon sources using Biolog plates, as well as custom plates incubated in environments containing different volatiles.

The strains chosen were a *Bosea* TC13b (97% identity based on 16S rRNA gene sequence), *Rhodococcus fascians* W9 (99% identity), *Methylobacterium* GM3 (94% identity), *Rhodococcus* bl28a (96% identity), and two *Rhodococcus*

erythropolis strains 1(2b)b and bl28ba (98%, 99% identity), each of which had been demonstrated to degrade isoprene added to liquid culture as determined by GC-FID analysis.

Each of the strains, apart from strain GM3, showed a wide range of metabolic potential, able to utilise numerous compounds, especially strain W9 (Table 3.SI.2). Strain GM3 had a much lower metabolic potential, with significant increases in respiration activity only with ribose, xylose, and lyxose carbon sources in the Biolog plates (Table 3.SI.2). With volatiles as carbon sources, interestingly, all strains had increased respiration in the presence of hexane, compared to with no volatile addition. Two of the strains shown increased respiration with toluene as well, and one, Tc13b even had increased respiration with DMSO. No strains significantly increased respiration with the alkanes, or benzene, and two did not significantly increase respiration when incubated with isoprene (despite being shown to in other experiments), possibly indicating a potential for false negatives due to low sensitivity.

Table 3.2 Significant increases in respiration for isolates with different carbon source additions.

Strain	1(2b)b	bl28a	bl28ba	GM3	Tc13b	W9
Species	*Rhodococcus erythropolis*	*Rhodococcus sp.*	*Rhodococccus erythropolis*	*Methylobacterium sp.*	*Bosea sp.*	*Rhodococcus facians*
Isoprene	0	**	*	*	*	0
Toluene	0	**	*	0	0	0
Hexane	*	**	**	*	*	**
DMSO	0	0	0	0	*	0
Propane	0	0	0	0	0	0
Methane	0	0	0	0	0	0
Ethane	0	0	0	0	0	0
Benzene	0	0	0	0	0	0

*** = p< 0.05, ** = p<0.01, 0 = no significant difference.**

3.4.4 Isoprene-degrading *Rhodococcus* have similar metabolic capabilities, but different genomic arrangements.

To investigate the isoprene degradation operon, and the metabolic capabilities further, the *Rhodococcus* strains 1(2b)b, bl28a and bl28ba were genome sequenced using PacBio technology and assembled using SMRT-Analysis software 2.3.0 (at NBAF-L) for analyses (see section 3.3.4 Genome sequencing for details). The *Rhodococcus* AD45 genomic data (van Hylchkama Vlieg et al., 2000) was also subjected to the same pipeline for comparison purposes.

Strain 1(2b)b had a genome size of 6.545 Mbp with one 0.4 Mbp plasmid, one 0.05 Mbp plasmid, and seven other contigs under 10 Kbp. The genome along with the large plasmid have high (99%, 98%) similarity to the respective genomic structures in *Rhodococcus* BG43 (plasmid pRLLBG43) (Ruckert et al., 2015) (with 91% and 16% coverage), confirming that the assemblies were likely complete. The small 0.05 Mbp plasmid was identified as the *R*. *erythropolis* NS1 pNSL1 plasmid (Valero-Rello et al., 2015). Strain 1(2b)b had a total size of 7.02 Mbp and a mol% GC content of 62.2%. RAST annotation of strain 1(2b)b lead to 442 subsystems (a biological process or pathway containing the predicted proteins), with subsystem coverage of 35%, where 2208 of 4407 genes not in subsystems were hypothetical (Figure 3.SI.5). Full KEGG metabolic pathway analysis is available in Figure 3.SI.6.

Strain bl28a had a genome size of 6.48 Mbp, a 0.45 Mbp plasmid, and 7 contigs under 10 Kbp, with a total size of 6.9 Mbp, and a GC content of 62.2%. RAST annotation indicated 2277 proteins in subsystems (100 hypothetical), and 4324 not in subsystems (2133 hypothetical) (Figure 3.SI.3). Full KEGG metabolic pathway analysis is available in Figure 3.SI.4.

Strain bl28ba had a total size of 6.9 Mbp and a GC content of 62.2% across 24 contigs, and reduction in contig number could not be achieved by further contig assembly in CAP (Contig Assembly Program) Hang (1999). RAST annotation led to 2283 predicted proteins in subsystems, of which 99 were hypothetical, and 4345 not in subsystems (2141 hypothetical) (Figure 3.SI.1). Full KEGG metabolic pathway analysis can be found in Figure 3.SI.2.

The genomes of the three strains were quite similar in terms of subsystem coverage and distribution of feature (defined DNA region, usually protein encoding, or RNA, prophage etc.) counts into subsystems, showing that these organisms have similar predicted metabolic capabilities.

Table 3.3 Key characteristics of genome sequenced strains

Strain	GC content	Total size (Mbp)	Largest contig size (Mpb)	Number of contigs >10Kbp
1(2b)b	62.20%	7	6.545	3
bl28a	62.00%	6.9	6.48	2
bl28ba	62.20%	6.9	5.9	4

3.4.5 The *Rhodococcus* AD45 isoprene degrading operon may be the exception

In order to investigate the isoprene degradation operon, the proteins sequences for IsoA-J (contributing to the isoprene degradation pathway) from *Rhodococcus* AD45 (van Hylckama Vlieg et al., 2000) were used as a t-BLASTn input, and searched against a custom BLAST database containing the genomes for each organism sequenced in this thesis. The location of the BLAST hits was used to identify and therefore allow for investigation of the isoprene degradation operon using RAST.

Once identified, the putative isoprene degrading operons from this thesis exhibited similar levels of dissimilarity to the isoprene degrading operon of *Rhodococcus* AD45 as each other, and a much more similar gene orientation between them than to *Rhodococcus AD45*: much more similar to that of alkB, and the *Gordonia* and *Mycobacterium* from Johnston (2014). However, despite the similarities between the isoprene degrading operons identified in this thesis and *Rhodococcus* AD45, in strains 1(2b)b and bl28a the operon is on the chromosome (~6.5 Mbp contig), whereas in *Rhodococcus* AD45, and strain bl28ba, the gene cluster is on a plasmid, suggesting that the plasmid association of the isoprene degrading operon is not ubiquitous (Figure 3.5, 3.6, 3.7, 3.8).

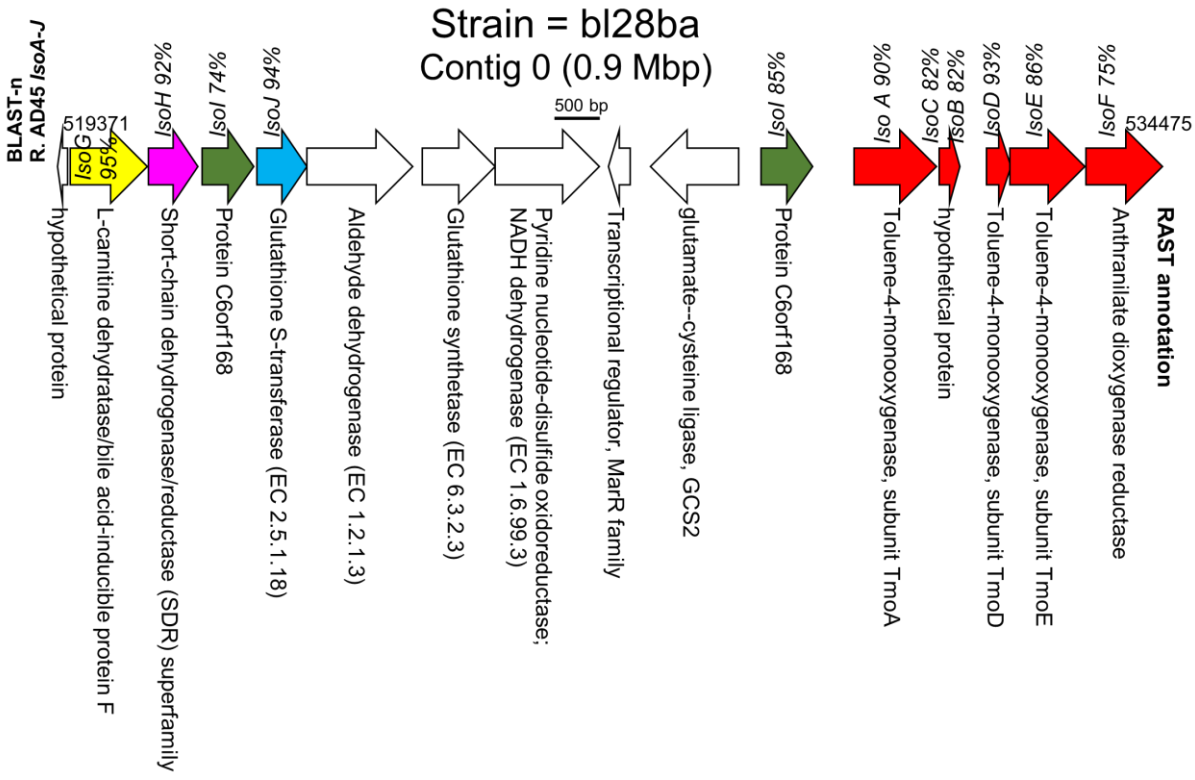


Figure 3.5 Genome representation of bl28ba cluster most similar to the *Rhodococcus* AD45 isoprene degrading cluster with above (anticlockwise turned) text representing the similarity to the equivalent AD45 protein, the below (clockwise turned) text displaying the RAST annotation, and the positional identifiers on either end. Colours are as with AD45.

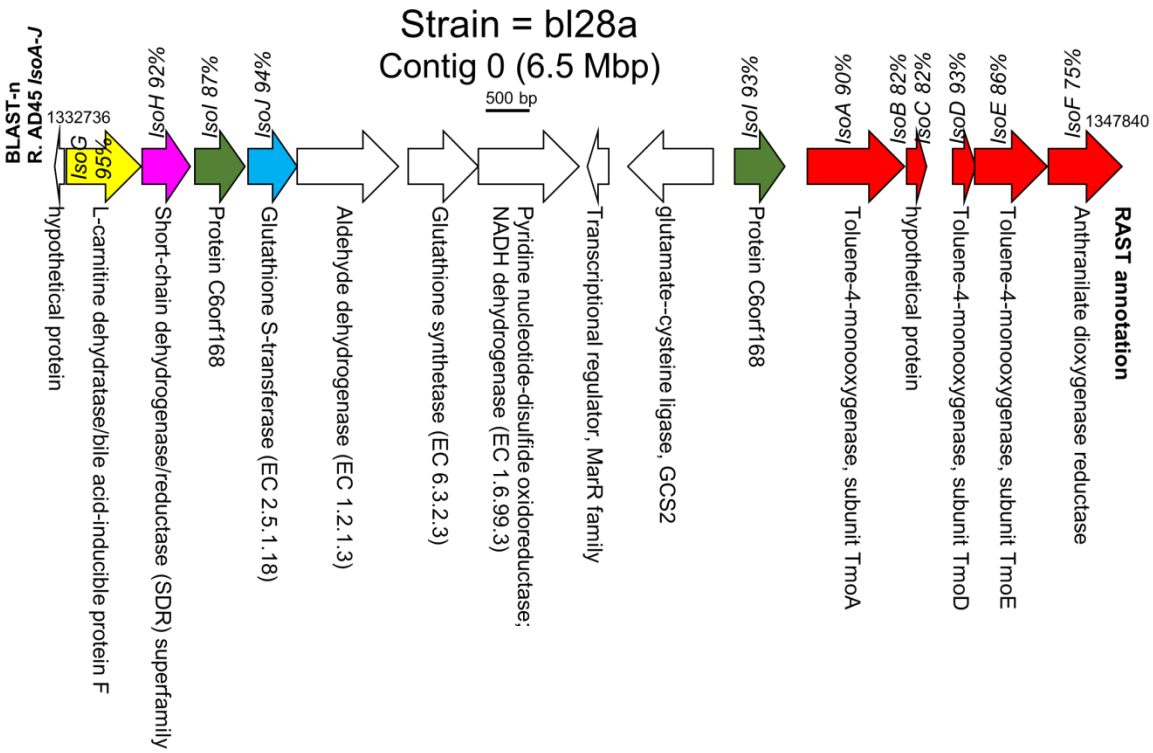


Figure 3.6 Genome representation of bl28a cluster most similar to the *Rhodococcus* AD45 isoprene degrading cluster with above (anticlockwise turned) text representing the similarity to the equivalent AD45 protein, the below (clockwise turned) text displaying the RAST annotation, and the positional identifiers on either end. Colours are as with AD45.

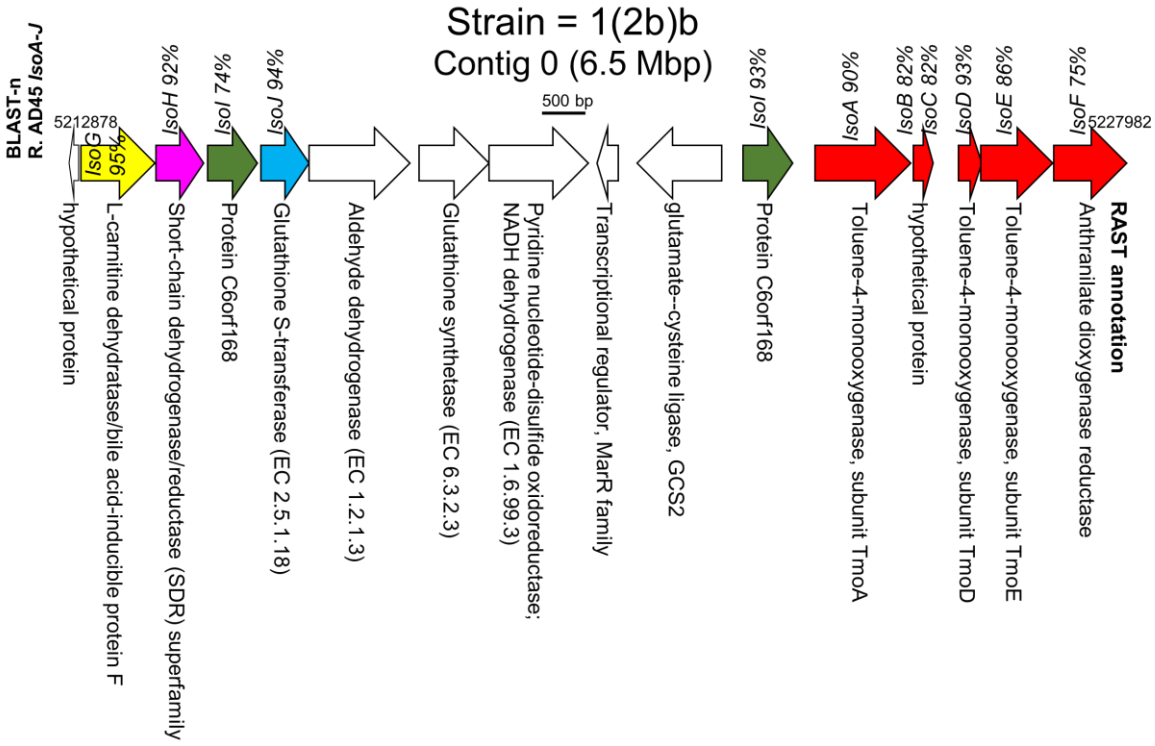


Figure 3.7 Genome representation of 1(2b)b cluster most similar to the *Rhodococcus* AD45 isoprene degrading cluster with above (anticlockwise turned) text representing the similarity to the equivalent AD45 protein, the below (clockwise turned) text displaying the RAST annotation, and the positional identifiers on either end. Colours are as with AD45.

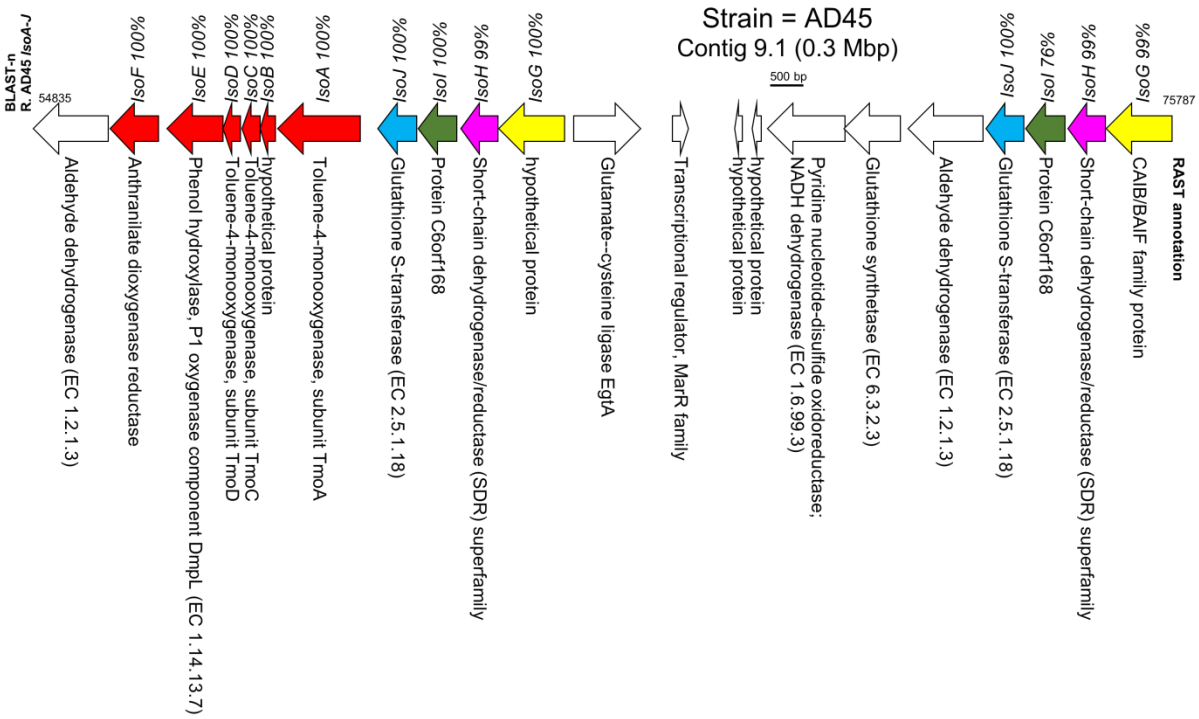


Figure 3.8 Genome representation of *Rhodococcus* AD45 cluster most similar to the *Rhodococcus* AD45 isoprene degrading cluster with above (anticlockwise turned) text representing the similarity to the equivalent AD45 protein, the below (clockwise turned) text displaying the RAST annotation, and the positional identifiers on either end.

3.4.6 Isoprene degrading genes from the strains in this thesis were similar to those from known isoprene degraders.

A t-Blastn analysis with *isoA* against a custom database containing the three genomes from this thesis and *Rhodococcus* AD45 showed that the putative isoprene degrading gene clusters the three strains in this thesis had around an 80% identity to *Rhodococcus* AD45, with no gaps. Short sequences in other regions than the putative isoprene degrading operon from bl28ba (×4) and 1(2b)b, as well as a short sequence from *Rhodococcus* AD45 also had partial matches (consistent with AD45 having a similar upstream region), suggesting similar motifs elsewhere. The strains also had a higher isoprene degrading

operon similarity to *Rhodococcus* PD630 than to *Rhodococcus* AD45 (not shown).

Searching the same database with t-BLASTn of the toluene monooxygenase system protein A from *Pseudomonas mendocina* KR1, from Uniprot (The UniProt Consortium, 2014) shown that the isoprene operon in the strains from this thesis had a lower similarity toluene monooxygenase than to isoprene monooxygenase. The toluene monooxygenase search found the *isoA* region, with AD45 at 48% identity, 1% gaps, bl28ba having 48% identity, 1% gaps, bl28a with 48% identity, 1% gaps, and 1(2b)b showing 48% identity, 1% gaps. The *Rhodococcus* AD45 isoprene degrading operon was more similar to the toluene monooxygenase from *Pseudomonas mendocina* KR1 than the equivalent regions in the strains from this thesis were.

As a control, propane monooxygenase/reductase mmoC (RHRU231_30052; A0A098BFN9_9NOCA) (Uniprot) from *R*. *ruber* gave less than 30% identity in all cases.

As all strains degrade hexane, AlkB, a hexane monooxygenase protein previously shown to be used in hexane degradation (as well as Butane, Octane and Decane), but with multiple copies (Hamamura et al., 2001), was used as a search query against the local genomic blast database; however no results had more than a 56% identity (and scores below 1215).

3.4.7 TM7 may be interacting with *Aeromicrobium*

In order to understand the reason why TM7 is enriched to a large extent in some isoprene enrichments, despite evidence for direct incorporation of carbon from isoprene not being found (Chapter 2), DNA from an enrichment containing 70.9% 16S rRNA sequences belonging toTM7 was extracted and sequenced thrice using Pacbio technology.

The metagenome resulted in 2.22 Mbp across 217 polished contigs, with a mean GC content of 59%, a mean sequence size of 10230, and a median of 9407, with contigs ranging from 506 bp to 37 Kbp. Additionally, two more sequence files were generated from the metagenome, one with TM7 identifiable contigs (22 contigs; between 2342 and 37234 bp, 315 Kbp total) and another with the metagenomic community without TM7 (sans-TM7 community). Less than 4.3% of the contigs were unidentifiable by blast, and 14.7% of the DNA was identified as deriving from TM7, which is expected as previously sequenced TM7 species have small genomes (~0.7 Mbp (He et al., 2015; Kantor et al., 2013)), this comprised of 0.324 Mbp spread across 23 contigs, representing between 1/3 and ½ of the estimated genome size.

The TM7 rich microbial community displayed a wide range of metabolic potential, and although the community as a whole lacked some of the metabolic functions of *Rhodococcus* alone (evidencing either the large metabolic diversity of *Rhodococcus*, or a limited diversity in the microcosm). Despite this, there were some metabolic pathways found in the community, but not in *Rhodococcus* (See Figure 3.SI.10 for full comparative KEGG metabolic analysis).

The TM7 only sequences were uploaded to RAST for annotation.

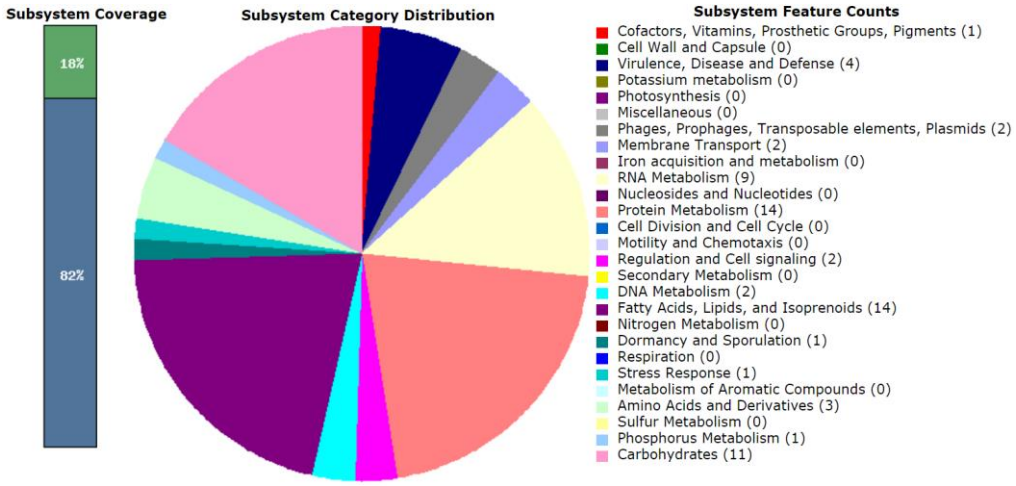


Figure 3.9 Predicted metabolic functions of TM7. Obtained by RAST annotation of TM7 sequences generated by filtering a TM7 rich microcosm for sequences identified as TM7 by their blast results. Numbers in brackets represent the number of proteins encoded in the sequence which fall into the subsystem (biochemical pathway).

Unfortunately, as TM7 has a small genome, the proportion of DNA returned as TM7 was much lower than that expected from its relative abundance as measured by amplicon sequencing. However, having an estimated 1/3 to 1/2 of the genetic information from non-overlapping contigs means that we should be able to get a reasonable idea of the metabolic capabilities. The annotation of the TM7 genome (Figure 3.9) showed predicted gene functions for: biotin biosynthesis, bile hydrolysis, beta-lactamase production,' a *Mycobacterium* virulence operon (a Jag protein, YidC and YidD), potassium metabolism, two phage tail proteins, two magnesium cation transporters, nine proteins involved in RNA metabolism, 14 involved in protein metabolism, an orphan regulatory protein, a stringent response protein, UvrD and Ycfh for DNA metabolism, 14 proteins for isoprenoid synthesis (six for carotenoids), one sporulation associated protein, a ROS protective protein, alanine, glycine and serine synthesis proteins, a phosphate metabolism protein, and 11 carbohydrate

metabolism proteins (one for pyruvate metabolism, five from the pentose phosphate pathway, two for trehalose synthesis, three for ribose utilisation).

In order to compare the total metagenomic community with the sans-TM7 metagenome, both were uploaded to MG-RAST, and annotated (Figures 3.SI.8, 3.SI.9). There were no factors likely to give TM7 an advantage identifiable through analysis of the difference between these data sets (ignoring carbon metabolism as isoprene the only major source present), as they had similar antibiotic resistances, and virulence factors.

Following this, the sans-TM7 community and the TM7 genomic data were compared in MG-RAST, and the areas where TM7 interacted with the metabolic capabilities of the rest of the metagenome were noted (Figure 3.SI.11). These enzymes at the terminal points were identified in *E. coli* and searched against a local BLAST database of the TM7 rich community. The results allowed identification of the contigs, and therefore the bacteria to which they belong could be identified through BLAST. TM7 gives to the community a number of capabilities including; the reversible reaction between Glycine + 5,10-Methylene-THF and Serine, folate / dihydrofolate interconversion, 5,6,7,8-Tetrahydrofolate (THF) production, an alternative route to make 5,10-Methylene-THF, Glycerate-1,3P2 and Glycerone-P production, Glycerone-P to Glyceraldehyde-P. Glycerate-3P (and output of glyoxylate and dicarbonate metabolism) to 1,3-Bisphosphoglycerate (and after that becomes 5,10-Methylene-THF), and the conversion of Glyceraldehyde-3P to Glycerone-P (Figure 3.10).

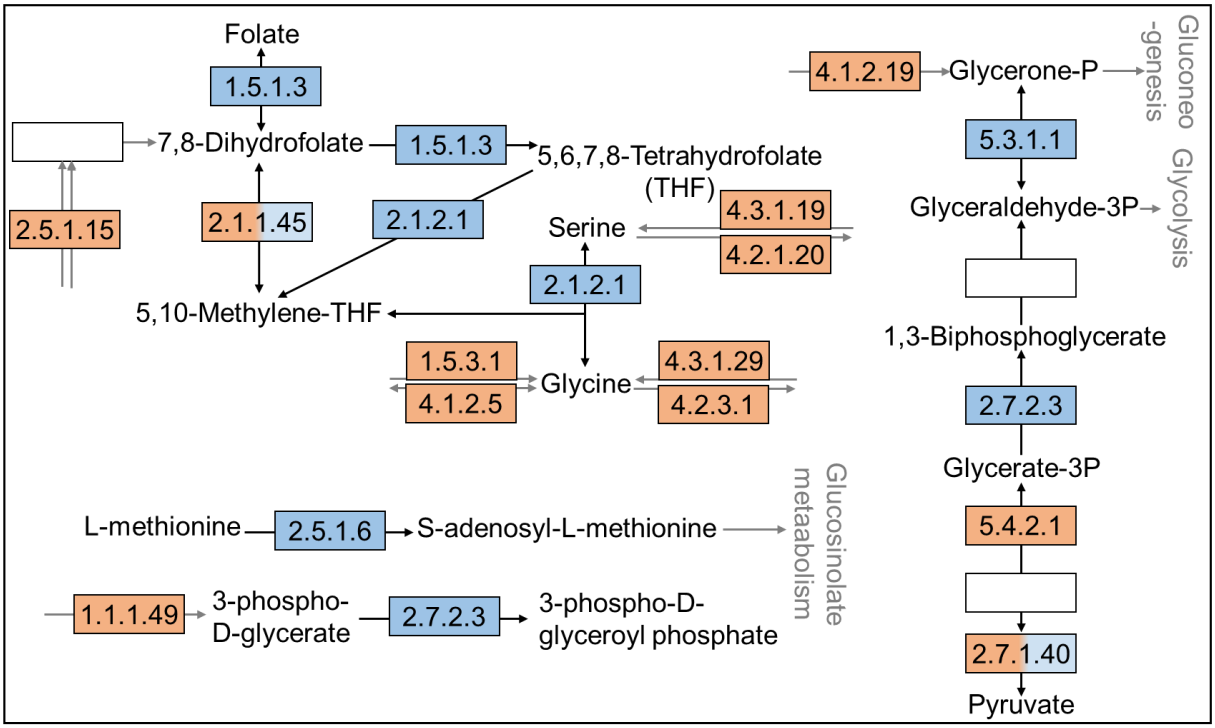


Figure 3.10 The proteins involved at the terminal and initial steps of TM7 metabolism; and the proteins from other groups which provide or utilise the product. Protein function prediction of a TM7-rich community was compared to that of the same community with TM7 sequences removed; with the difference being contributed to the metabolic potential by TM7. Blue boxes are TM7 proteins, white is where no hits are in the local BLAST database, orange is where other bacterial enzymes are predicted to interact with the substrate/product from the metagenome.

tBLASTn of the Kegg identified E. coli representative amino acid sequence for the proteins led to: 2.5.1.15 being identified as belonging to *Clavibacter* or *Leifsonia*, 1.5.3.1 as belonging to *Rhodococcus* or *Dermacoccus*, 4.1.2.5, as belonging to *Aeromicrobium*, 4.3.1.29 as belonging to *Pedobacter*, 4.3.1.19 as belonging to *Amacolatopsis*, 4.2.1.20 as belonging to *Aeromicrobium*, 5.4.2.1 as belonging to *Aeromicrobium*, 1.1.1.49 as belonging to *Aeromicrobium*, 2.1.1.45 (both) as belonging to *Aeromicrobium*, and 2.7.1.40 (both) as belonging to *Aeromicrobium* (4.2.3.1, 4.1.2.19 were not found).

3.5 Discussion

3.5.1 No phylogenetic linkage observable between *isoA* gene sequence and *16S rRNA* gene sequence

From the analysis of *16S rRNA* and *isoA* genes for the collection of bacteria isolated as part of this thesis, there is a greater diversity of isoprene-degrading bacteria than previously thought, including the first isolates/examples of Phyla such as *Pimleobacter*, *Pedobacter*, *Bosea*, *Stenotrophomonas*, *Arthrobacter*, *Nitrobacter*, *Ensifer*, *Variovorax*, *Devosia*, and *Methylobacterium.* This suggests that there is widespread capacity for isoprene degradation amoungst bacteria, and provides a resource for further study.

A study based on n-hexadecane degradation has demonstrated that a number of those genera have representatives capable of degradeing other hydrocarbons, including *Nitrobacter*, *Bosea*, *Variovorax*, *Stenotrophomonas*, and *Pedobacter*, which the authors also note is likely based on the presense of the *alkB* gene, which is mobile, and has no link between phylogeny and *alkB* sequence (Giebler et al., 2013).

Likewise, when (in this thesis) *isoA* gene sequences and *16S rRNA* gene sequences from strains shown to degrade isoprene were compared to the other strains, there was no presence of *isoA* phylogenetic linkage, the *isoA* gene sequence similarity to other *isoA* sequences did not correlate with the *16S rRNA* gene sequence similarity (of the same bacteria) to other *16S rRNA* gene sequences).

However, the increased *isoA* sequence dissimilarity visible in Figure 3.3 with the strains from this project (which are in clusters 1,2 and the (comparatively) highly diverse cluster 3, whereas strains from El Khawand (2014) are in cluster 1, and the cluster 1-2 border, only) does raise the question of whether the

increased diversity is actually true *isoA* diversity, or if the primers were not being sufficiently specific to pick up *isoA* gene sequences only. It is possible that the touch-down PCR conditions with a high total number of cycles developed for these primers, with a large number of cycles annealing at (the quite low) 54°C, although within range for the primer predicted Tm, could increase chance of mis-priming and incorrect product, or even contamination thresholds. El Khawand (2014) when designing the primers did note a non-specific product at 54°C annealing temperatures, and implemented the hot start protocol to reduce this (this does not eliminate this). El Khawand tested the primers against high isoprene enrichment cultures and against strains with a several monooxygenases (including a toluene monooxygenase with no visible product). The sequences from the enrichments were all greater than 96% similar to *Rhodococcus* AD45 (although from the *isoA* clones you cannot tell if they some were from extremely similar organisms), and although the *isoA* genes from the monooxygenase amplification were amplified to a higher concentration, there was a product of the right size with the methane monooxygenase sample from a *Methylococcus capsulatus*, and a shorter product from a *Methylocella silvestris*, containing a Methane monooxygenase and propane monooxygenase; showing that the primers were not entirely specific).

Although, considering that the product was confirmed to be visible and was the correct size in each case, and considering that the isolates grow on isoprene, it is likely that we are looking at a diverse set of gene sequences involved in isoprene degradation (although possibly opportunistically given some of the

high concentrations of isoprene used in the enrichments), even if not necessarily isoprene specific monooxygenases.

3.5.2 A different, more diverse operon involved in isoprene degradation exists

The genomic data supports the idea of isoprene degradation not being strongly phylogenetically linked, with isoprene degrading operons being more similar in some diverse bacteria (e.g. the *Gordonia*, and *Mycobacterium* strains from Johnston 2014) to AD45, than others identified in this project. In two out of three genomes sequences from this thesis, the isoprene degrading genes are on the main chromosome, which, considering the likelihood of horizontal transfer (a likely cause of phylogenetic gene sequence dislinkage) and the presence of the isoprene degrading operon in the plasmid in some isolates, this suggests possible episomal behaviour of the isoprene degrading operon.
It is worth noting when considering the results from this thesis, and other work that should horizontal gene transfer of the isoprene degrading operon be an important factor, strains able to grow on isoprene, may have gained that ability whilst in being enriched – and therefore are possibly not reflective of the isoprene degraders in the environment.

3.5.3 TM7 may be interacting with *Aeromicrobium*

The interaction between TM7 and the other bacterial strains is interesting, or it is potentially coincidental that most of the genes in TM7 which are also absent in the rest of the metagenome involve parts of the glycolysis/gluconeogenesis

pathways, or THF involving steps, hinting at interactions with folate and glucose metabolism.

Considering that all of the non-redundant metabolic interactions with TM7 seem to be with *Aeromicrobium*, this indicates that *Aeromicrobium* may be the genus that contains what TM7 requires to thrive, potentially this being Glycerate-3P, and 3-phospho-D-Glycerate, Interestingly glyceraldehyde-3-phosphate normally is from photosynthesis, but here is being provided by *Aeromicrobium*, and is also present at the start of the MEP pathway of isoprenoid biosynthesis, suggesting at a more complex interaction. This could make *Aeromicrobium* an interesting candidate for attempting TM7 co-culture. Increased sequencing depth (to get the whole genome of both, and the community) would give much more strength to these results.

3.6 Conclusion

To conclude; it is certainly likely that the previously identified isoprene degrading operon is not the only operon involved in isoprene degradation. This is supported by (i) the phylogenetic dislinkage of *isoA* diversity, (ii) the comparatively large difference between the AD45-like *isoA* DNA sequences, and the bl28a-like *isoA* DNA sequences causing clear grouping into clusters, and (iii) the variable genomic location of the isoprene degrading operon. This may be an alternate system, or possibly the results here, previously, or both, are being affected by the high concentration. It is also worth noting that the orientation of the genes in the operon in *R.* AD45, with a slight double operon, when compared to the orientations of the genes in the isoprene degrading

operons in this study, and in *Mycobacterium* and *Gordonia,* and other monooxygenases clusters is likely the exception rather than the rule.

Additionally, this work highlighted that TM7 may be metabolically interacting with *Aeromicrobium* species, with some other possible interactions, although as this is based on predicted functions on partial genomes from metagenomic data, it is only indicative. Although this is backed up by a high level of co-occurrence of *Aeromicrobium* with TM7 from the sequential enrichment data in chapter 2, with co-occurrence in 13 of 19 low concentration microcosms, and 70 out of 71 serial enrichments, the concentration is low (~2%) in the later serial enrichments of Willow soil, and partially negatively correlated, so the cause of TM7 enrichment there may be different, or may involve the death of *Aeromicrobium*.

Future work on this topic should probably involve investigation of this alternative isoprene degradation operon, including using gene knockouts, should include alternative carbon source testing of *Rhodococcus* AD45-like strains, and should look at the TM7 interactions with more sequencing depth. Additionally, it may be interesting to combine *isoA* primers with SIP and next generation sequencing to increase depth, diversity and evidence isoprene degradation in bacteria which have the gene.

Chapter 4: isoprene degradation on the leaf surface

4.1 Abstract:

Isoprene is one of the major volatiles emitted from plants, which globally produce over 580 million tons per year. Most isoprene is attenuated by reactions in the atmosphere, however some microbes have been shown to be able to degrade isoprene in soil and in culture. As isoprene is released mainly from the leaves of plants, there is the likelihood that bacteria may be utilising this at the interface of the leaf surface and the atmosphere, reducing the net production and influencing the bacterial community. To investigate this, I investigated the effect of plant isoprene production on bacterial communities, and the effect of isoprene degrading bacterial communities on plant isoprene production, through: (i) incubating leaves in the presence of isoprene and measuring changes in isoprene concentration, (ii) investigating community changes by sequencing *16S rRNA* genes of bacterial communities from heterozygous isoprene-producing and azygous non-producing tobacco, (iii) inoculating leaves with isoprene degrading bacteria to determine whether there was a decrease in net isoprene production, (iv) creating isoprene degrading enrichments to investigate isoprene degradation rates and to isolate isoprene degrading microorganisms. There were no significant differences in bacterial communities between the isoprene producing and non-producing tobacco, nor was there significant isoprene degradation by on-leaf bacteria, however TM7 was on 4 out of 9 isoprene producing tobacco leaves, and was not detected on any non-producing tobacco leaves. Inoculation of tree leaves with isoprene

degrading bacteria generated no difference in net production. Isoprene degradation by leaf derived microbes in enrichments was observed in 31 out of 48 enrichments, resulting in a wide range of isolates, but was highly stochastic. The lack of significant differences suggests that, although isoprene degrading bacteria do live on the leaf surface, it may not be a major carbon source, and therefore its presence does not reproducibly influence the bacterial community; nor does inoculation of isoprene producing leaves with non-native isoprene degrading bacteria alter the net production rate after six weeks.

4.2 Introduction:

Isoprene (2-methyl-1,3-butadiene) is a highly reactive hydrocarbon that can have varying consequences depending on the prevailing atmospheric conditions. Isoprene can react with: (i) free radicals and indirectly increase the residence time of methane (Collins et al., 2002), (ii) nitrous oxides and create tropospheric ozone (Jenkin & Clemitshaw, 2002), and (iii) numerous other compounds creating various byproducts, including carbon monoxide (Pfister et al., 2008), organic aerosols (Paulot et al., 2009), and formaldehyde (Pfister et al., 2008). As methane contributes to global warming, and trophospheric ozone is harmful to plants and respiratory systems; isoprene can have negative effects on health, the climate and the economy (Ashworth et al., 2013).

Global production of isoprene is approximately six hundred million tonnes annually, with broadleaf trees responsible for 51%, and most of the rest (46%) coming from shrubs (Guenther et al., 2006). It is believed that isoprene is produced as a response to transient heat stress, likely crossing and stabilising

thylakoid membranes in leaves as sunlight causes temperature fluctuations of up to 10°C higher than the surrounding air (Sharkey et al., 2008; Singsaas et al., 1999).

Most atmospheric isoprene is attenuated through chemical reactions, which, due to the presence of two carbon-carbon double bonds, can be very rapid (Guenther et al., 1995), leading to an atmospheric lifespan of around 1.7 hours (Jenkin & Clemitshaw, 2002), and therefore a comparitively low atmospheric concentration, typically between 1 ppt and 10 ppb, depending on location (Baldocchi et al., 1995; Kesselmeier et al., 2000; Lewis et al., 1997). Despite this, there is evidence of biological isoprene degradation: isoprene has been shown to be biologically removed in soils, with an estimated global removal of 20 Tg per year (Cleveland & Yavitt, 1997) and isoprene-degrading bacteria have been isolated, and a probable pathway found (Acuña Alvarez et al., 2009; van Ginkel et al., 1987; van Hylckama Vlieg et al., 2000; Crombie et al., 2015)

In other hydrocarbon cycles, the hydrocarbon is often consumed proximate to its source; for example, low affinity *Methanotrophs* living in the oxic zone between the anoxic methane producing soil and the atmosphere degrade 90% of methane before it reaches the atmosphere (King, 1992; Conrad, 2009). For this reason, it is possible that a significant proportion of isoprene could be degraded by leaf dwelling microbiota before entering the atmosphere.

Leaves are inhospitable to most bacteria due to having low nutrient availability, the threat of physical removal by rainwater or wind, exposure to harsh

conditions due to fluctuations in temperature, humidity, water activity and solar radiation, often coupled with a yearly habitat destruction and renewal due to leaf loss. However, despite this seemingly bleak outlook for phyllosphere dwelling microbes, there is an average of 10^6-10^7 bacteria per cm^2 leaf surface, totalling 10^{26} bacteria globally (Lindow & Brandl, 2003), and the sometimes low level of available carbon sources on leaves drives epiphytic (and endophytic) bacterial coexistence (Wilson & Lindow, 1994), and could theoretically make isoprene a valuable carbon source.

The phyllosphere is generally a hostile environment for microbiota, and although Sucrose, fructose and glucose are detectable on the leaf surface, it is thought that other carbon sources, such as VOCs are generally not available to epiphytic microbiota (unless leaching, wounding, excretion, exudation or infiltration occur) due to the waxy cuticle (Nadalig et al. 2011; Vorholt 2012). This waxy cuticle can have large effects on the community composition. For example, there are large differences in community structure between different *Aribadopsis thialiana* plants with different mutations involved in the wax biochemical pathways, likely due to the changes in cuticle structure (Bringel and Coulee 2015).

Phyllosphere microbiota have been shown to degrade other hydrocarbon based products. Due to low levels of fungi and archea on the leaf surface, Bacteria are thought to be responsible for most of this degradation of hydrocarbons (Vorholt 2012). Some of the potentially important hydrocarbon degrading phyllosphere dwelling bacteria include *Phascolus vulgaris*, which has been shown to have *cph* genes involved in 4-chlorophenol degradation induced on phyllosphere, numerous *Methylotropic* bacteria which are ubiquitous on plants, and can use

plant compounds in situ, such as methanol, formaldehyde and chloromethane. Additionally, presence of Candida biodinii presence reduces the net emissions of methanol from Nicotiana seedlings, suggesting some fungal involvement (Bringel and Coulee 2015).

Of particular interest is microbiological degradation of chloromethane. Chloromethane is the most abundant VOC in the atmosphere (600 ppt).

Chloromethane degrading bacteria have been shown to exist on the leaf surface of *Aribadopsis thialiana* plants, and have been isolated with chloromethane as a sole carbon source (on minimal media plates, following minimal media solution enrichment of *A. thialiana* leaves). This resulted in four *Hyphomicrobium* species isolated. The gene organisation for the chloromethane degrading pathway was similar in these organisms, as well as to previously characterised strains (*Roseovarius*, *Rhodobacteraceae*, *Aminobacter*, *Hyphomicrobium*, *Methylobacterium* and Pseudomonas species), with the exception of *Metthylobacterium exotorquens* CM4 (Nadalig et al. 2011).

Through these, and other potentially methods, it has been suggested that phyllosphere based phylloremediation is possible. Pollutants could be removed from the atmosphere which are generated from external sources, such as nicotin, phenol, and polycyclic aromatic hydrocarbons, or from the plant, such as chloromethane and isoprene (Bringel and Coulee 2015). Potential candidates for this include species of *Arthrobacter*, which has been shown to degrade a number of aromatic hydrocarbons and are able to grow and remain in the phyllosphere.(Bringel and Coulee 2015).

The aim of this study is to determine whether isoprene degradation is happening on the leaf surface, and conversely, if isoprene production is capable of sustaining, or altering the communities of plant leaves; and if altering bacterial communities of isoprene-producing leaves by adding isoprene degrading microbes can reduce the net isoprene release of leaves, providing a biotechnological solution to isoprene emissions from isoprene producing biomass crops. To achieve these aims, the objectives were to incubate leaves with isoprene, measuring changes in the isoprene concentration, to grow isoprene producing and non-producing tobacco and investigate the bacterial community, and to perform isoprene enrichment of microbiota from leaves, obtain isolates and measure isoprene degradation rates.

4.3 Materials and Methods

4.3.1 Experimental Design: Tobacco communities.

Genetically modified *Nicotiana tabacum* (Tobacco) Line 32, both azygous isoprene-non-producing and heterozygous isoprene-producing, containing a *Populus alba* isoprene synthase gene (Vickers et al. 2009), were obtained from Claudia Vickers. Seeds were planted in Levington Seed & Modular Compost + Sand, watered with Hoagland's nutrient solution (Hoagland and Arnon, 1949), and incubated in the 20°C controlled greenhouse at the University of Essex. Two-week-old 2-3 cm seedlings were transplanted into discrete pots, nine recombinant isoprene-producing plants and nine control plants, and separated into trays of three with an alternating configuration (Figure 4.1). Row locations were switched at four weeks. At six weeks, when plants were about 20 cm tall, three leaves were sampled from each plant, maximising leaf distance and avoiding those either in direct contact or immediately neighbouring sampled leaves. Tobacco plants were healthy with no obvious differences between the groups. Leaves were cut into ~2 cm^2 sections and suspended in 20 mM Tris, 10 mM EDTA, 0.024% Triton X-100, sonicated for ten minutes and vortexed for one minute, before filtering through sterile glass wool to remove coarse plant material.

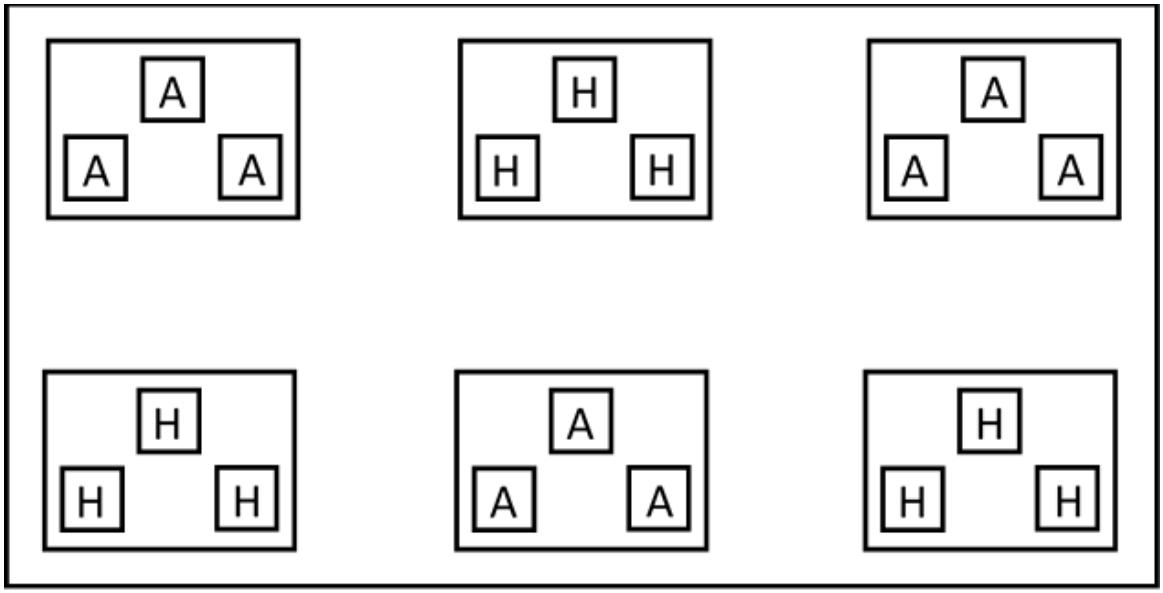


Figure 4.1 Initial isoprene-producing tobacco experimental layout; Tobacco plants were laid out in three blocks of three for each tobacco type, with alternating blocks .'H' indicates heterozygous isoprene-producing tobacco; 'A' represents Azygous, tobacco.

DNA was extracted from the cells derived from the tobacco leaf washings using the method described in Griffiths et al. (2000), and prepared for MiSeq by PCR as described in Chapter 2 (using the Illumina protocol (Illumina Inc., 2013), using RedTaq instead of Kapa HiFi, and adding one more cycle for the first round of PCR). MiSeq was performed, by the NBAF facility in the Centre of Genomic Research, Liverpool, UK.

4.3.2 Bioinformatics: Tobacco bacterial communities

Sequences were error corrected using the BayesHammer error correction implementation in SPAdes (Bankevich et al., 2012). Error corrected sequences were trimmed to a quality value limit of '18' using Sickle (Joshi & Fass, 2011), and paired using PandaSeq (Masella et al., 2012), before concatenation of the different runs and barcoding. Paired sequences were then de-replicated using the FastX collapser from the FASTX-Toolkit. Additionally, OTU binning was

carried out using Swarm (Mahé et al., 2014). Chimera checking was performed using the chimera checking option of vsearch (Rognes et al., 2015) against the Gold database (Pagani et al., 2012). All bioinformatic analyses were carried out within Biolinux 8, using default parameters if not specified otherwise. Swarm (d=1) and vsearch were carried out on the University of Essex Genomic Cluster (Debian OS). OTU table formation, non-chimeric sequence fetching and rarefication using true random numbers (Haar, 2008) were performed using custom algorithms implemented in Microsoft Excel 2013 using VBA. OTUs were classified using the RDP OTU classifier with the training set '16S rRNA 14' (Wang et al., 2001). Sequences classified as chloroplast sequences were removed and sequences which the RDP classifier assigned as "*Rhodoligotrophos*", which after nblast-ing (Altschul et al., 1990) turned out to be tobacco mitrochondrial DNA, were also removed. Remaining sequences were converted to relative abundances. Significance testing was performed through multivariate analysis tools for abundance data within the package mvabund in R. The manyglm function was used to fit negative binomial models, with *P* values calculated using 1000 resampling iterations (PIT trap resampling) (Wang et al., 2012).

4.3.3 Enrichment of leaf associated bacteria with isoprene.

Leaves from willow, oak and poplar were sampled from Wivenhoe Park, UK (tree details are in Table 4.1). Several leaves of varying sizes were sampled from 1-2 m in height from all around each tree. Leaves were cut into 2 cm^2 sections (c. 150 cm^2 leaf area per tube), sonicated for 10 minutes and vortexed for one minute whilst immersed in 30 ml 20 mM Tris, 10 mM EDTA, 0.024% Triton X-100 leaf wash solution, and filtered through sterile glass wool (Ikeda et

al., 2009). Leaf washings (0.1 ml) from willow, oak, poplar (or tobacco in a separate experiment) or 0.1 ml sterile water (as a control), was added to 9.9 ml of minimal media with the composition; 0.5 g NaCl, 0.5 g $MgSO_4 \cdot 7H_2O$, 0.1 g $CaCl_2 \cdot H_2O$, 1 g NH_4NO_3, 1.1 g Na_2HPO_4, 0.25 g KH_2PO_4, 50 mg Cycloheximide, 10 mg $FeSO_4.7H_2O$, 0.64 mg $Na_2EDTA.{}_3H_2O$, 0.1 mg $ZnCl_2$, 0.015 mg H_3BO_3, 0.175 mg $CoCl_2 \cdot 6H_2O$, 0.15 mg $Na_2MoO_4 \cdot 2H_2O$, 0.02 mg $MnCl_2 \cdot 4H_2O$, 0.01 mg $NiCl_2 \cdot 6H_2O$, 0.05 mg *p*-Aminobenzoic acid, 0.02 mg Folic acid, 0.02 mg Biotin, 0.05 mg Nicotinic acid, 0.05 mg Calcium pantothenate, 0.05 mg Riboflavin, 0.05 mg Thiamine HCl, 0.1 mg Pyridoxine HCl, 0.001 mg Cyanocobalamin, and 0.05 mg Thioctic acid (phosphates and salts were autoclaved separately, and vitamins were filter sterilised) per litre Milli-Q filtered water (a modification of Fahy et al., 2006), in a 125 ml glass serum bottle sealed with a PTFE backed silicon septa. To each serum bottle, for each leaf type, either 0.01% yeast extract, 5 g of 2 mm diameter sterile glass beads, or a small leaf (from the same tree) were added in triplicates, along with a triplicate with no additional substrates or surfaces. To this, 0.1 cm^3 saturated 30°C isoprene headspace was added, yielding a concentration of 7.2×10^5 ppb isoprene. The isoprene concentration in the enrichment was monitored daily using 100 µl headspace sampling onto a Unicam 610 Gas Chromatograph with a 10% Apiezon L CW column and a flame ionisation detector (GC-FID), with injector and detector temperatures of 160°C and a column temperature of 100°C. Enrichments showing greater than 70% isoprene degradation after 15 days were diluted 10^1 to 10^4 fold and samples were spread onto minimal media plates of the same composition with agar added to 1.5%. Representative colonies were picked and re-plated.

Table 4.1 Tree location for Wivenhoe Park samples

Common name	Species	Location °N	Location °E	isoprene producing
Willow	*Salix babylonica*	51.877720	0.948709	+
Birch	*Betula pendula*	51.878280	0.950293	-
Oak	*Quercus robur*	51.877990	0.949446	+
Poplar	*Populus canadensis*	51.878151	0.948988	+

4.3.4 Experiment to determine isoprene-degradation on detached oak leaves

Eighteen oak leaves were collected on 1st June and 28th September 2015 from Wivenhoe Park UK, and three leaves were placed into each leaf chambers comprising of pouches constructed of 150 µm Polyethylene Terephthalate layer with an ethyene-vinyl acetate adhesive coating, cut to a size of 105 × 148 mm, containing small (25 mm) computer style fans powered by an external 12 V battery, and PTFE/Si septa held by Swagelok 316 stainless steel 1/16" bulkhead unions (Swagelok) (Figure 4.AX.1). Chambers were wrapped in aluminium foil to prevent light entry. Control chambers without leaves were set up at the same time.

Figure 4.2 Photograph of leaf chambers, with leaves, without foil coverings

Each leaf chamber was heat sealed on the opening tabs (ensuring no contact with the leaf) within 30 minutes of sampling, allowed to acclimatise for 15 minutes, evacuated and refilled with 50 cm^3 6 ppm isoprene and 6 ppm of an inert tracer, 1,1,1,2,3,4,4,5,5,5-Decafluoropentane (DFP), immediately prior to the time zero (T0) measurement of isoprene and DFP concentration using the GC-FID. Further headspace GC-FID sampling was carried out after 20, 80 and 320 minutes. Isoprene concentrations were adjusted for any leakage by using the isoprene to DFP ratio.

4.3.5 Experiment to test the effect on net isoprene production by leaves after inoculation with isoprene-degrading bacteria

The leaves of three small branches of willow, ash and birch trees were inoculated with isoprene-degrading bacteria. A culture containing a mix of bacteria capable of rapid isoprene degradation, was created by sequential

enrichment of soil that originated from within one metre of the base of the Willow tree in Table 4.1. One gram of sample was suspended in 9 ml of minimal medium, and incubated with 7.2×10^5 ppb isoprene in the same manner as the leaf microcosms. Following greater than 80% degradation, 1 ml of the enrichment was inoculated into new media, and repeated for eight further generations. The final enrichment was dominated by *TM7* (30%) and *Rhodococcus* (23%), with the rest mainly consisting of *Pelomonas, Leifsonia, Acinetobacter, Pseudomonas and Stenotrophomonas* based on amplicon sequencing (see chapter 2). On the 13th August 2015 three test and three control branchlets, each between 0.5 and 2 m above the ground, and over 90° angle of each other around the trunk, on the willow, oak and birch trees (referenced in Table 4.1) were thoroughly sprayed with either the microcosm diluted to OD 0.45 (measured on a spectrophotometer) with minimal media, or with minimal media alone. On the 28th September 2015, leaves were collected into non-reactive leaf chambers with air agitation, sealed (dark, STP), and isoprene production was monitored by GC-FID.

4.4 Results

4.4.1 Tobacco leaf community structure is not significantly driven by isoprene production

Transgenic heterozygous isoprene-producing tobacco and azygous non-producing tobacco were grown for six weeks in three alternating groups of three, followed by sampling of leaf microbiota and amplicon sequencing (as above).

No significant differences in composition was found between the bacterial communities from the isoprene-producing and the non-producing tobacco either when the tobacco plants were treated as individual replicates, or as replicates in their blocks for any of the bacteria sequenced (Anova), nor were there differences at the community level between the groups (Permanova $F_{1,16}$ =1.327 P=0.208) (Figure 4.2).

The bacterial communities from both the isoprene-producing and non-producing plants were dominated by *Rhodococcus*, contributing approximately 54% of the sequences. In the communities from isoprene-producing tobacco, there were increased *Rhodococcus* (from 34% r=to 48%) and *Pelomonas* (from 22% to 26%) compared to the non-producing tobacco communities, and candidate division *TM7* (9% average) was detected in four of nine replicates from isoprene-producing tobacco and not detected in any of the non-producing tobacco replicates, although none of these differences were significant (p>0.05). TM7 was not evenly distributed in the plants on which it was detected, with relative abundances of 70%, 7.4%, 2.0%, and 0.8% for the four leaves it was present on.

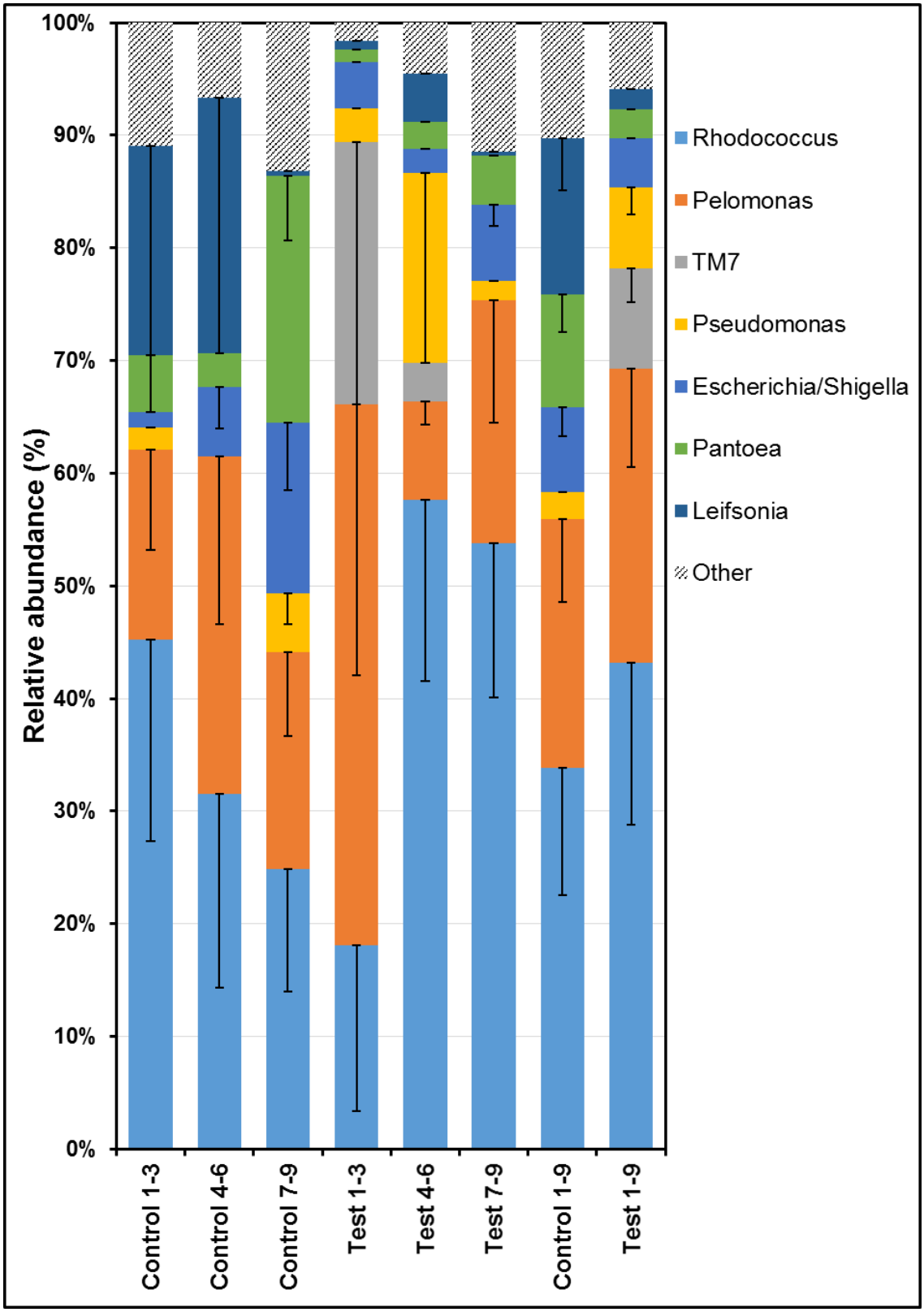


Figure 4.3 Bacterial leaf communities for isoprene-producing and non-producing tobacco after six weeks plant growth. Test represents community from heterozygous isoprene-producing tobacco, Control is the azygous isoprene-non-producing tobacco community. Error bars are negative standard error. n = 3, n = 9 for Control 1-9 and Test 1-9. Strains with an all plant average less that 5% are grouped as "Other".

4.4.2 Degradation of isoprene by on-leaf microbiota is undetectable

Oak leaves were sampled in June and September 2015. Changes in isoprene concentration were recorded and compared to that of the control to determine potential isoprene degradation.

Initially (from 0 to 20 min), the isoprene concentrations decreased significantly from oak leaves sampled in both June and September (Anova $p<0.05$), which equates to isoprene consumption of around 15.7 ppb min^{-1} cm^{-2} (leaf) cm^3 air . However, from 80 min any decrease was overwhelmed by an increasing isoprene concentration (Figure 4.4).

.

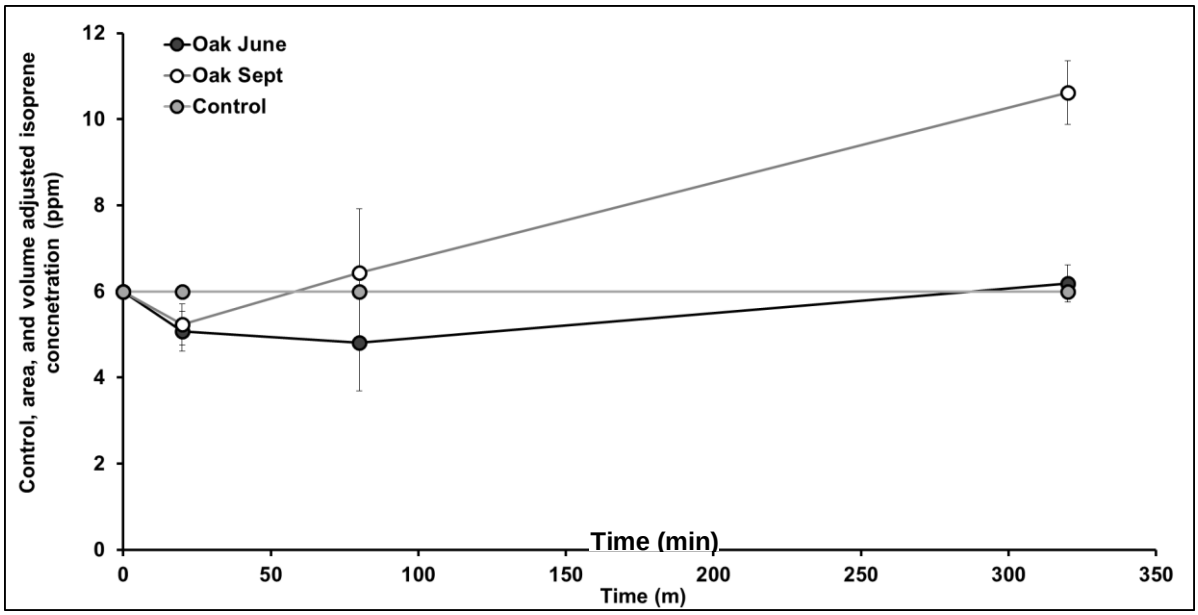


Figure 4.4 Changes in isoprene concentration on detached oak leaves adjusted for escape against decafluoropentane concentrations compared to no-leaf control. Control-adjusted values, Error bars are Standard Error. n=3.

4.4.3 The addition of isoprene-degrading bacteria to tree leaves did not lead to a long term difference in net isoprene production

A mixed culture, consisting of mainly *TM7* and *Rhodococcus* species, was sprayed onto leaves of Oak, Willow and Birch, and compared with media only controls. Leaves remained attached to the trees and were sampled after 6 weeks, when they were sealed into leaf chambers. There was no significant difference in isoprene production between the test and control groups (Figure 4.5).

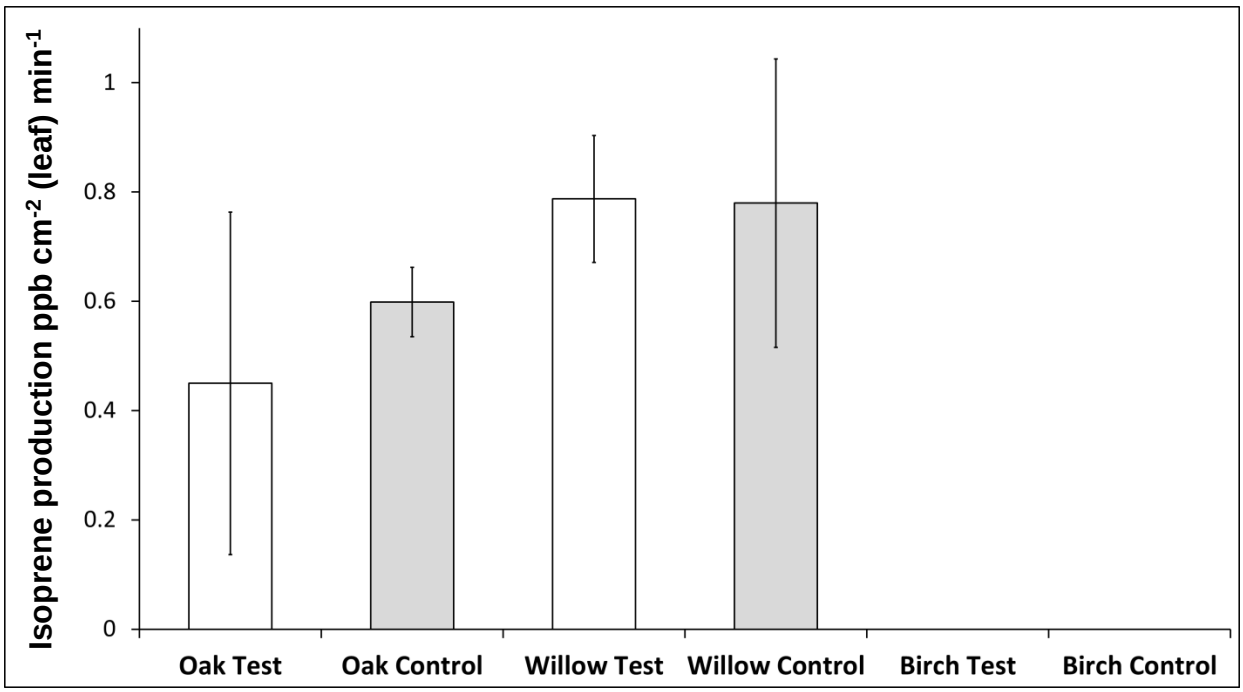


Figure 4.5 Net production of isoprene with treated and un treated leaves following a six week settling period, after three hours of confinement in 50 cm^3 leaf chambers. n=3.

4.4.4 Leaf bacteria degraded isoprene, and isoprene degrading bacteria are present on a wide variety of leaves

Leaf washings, from poplar, willow and oak trees and from heterozygous isoprene-producing and azygous non-producing tobacco, were incubated in minimal media with isoprene; and isoprene concentrations were monitored by GC-FID.

Cells detached from tree leaves degraded isoprene, demonstrating microbial degradation of isoprene by leaf dwelling microbes. However, the isoprene degradation was highly variable, with only 25% of isoprene-enrichments degrading more than 60% of the isoprene after 15 days, and 35% of the isoprene-enrichments showing no or minimal (less than 30%) degradation (Table 4.2).

Table 4.2 Isoprene enrichments of cells detached from tree leaves showing a decrease in isoprene after 15 days.

% Degradation	Total number of samples	Willow	Poplar	Oak
<30	17	6	5	6
30	10	2	4	4
40	5	3	0	2
50	4	2	2	0
60	4	2	0	2
70	4	0	2	2
80	2	0	2	0
90	2	1	1	0
Total:	48	16	16	16

There was no degradation of isoprene after over 200 days in leaf wash enrichments from isoprene-producing and non-producing tobacco (Figure 4.6). One of the 3 replicate enrichments inoculated with microbial cells from isoprene-producing tobacco did show a decrease in isoprene. As the other

replicates remained very close to the control group, it was possibly a leak, and the difference between the means for each group was not significant.

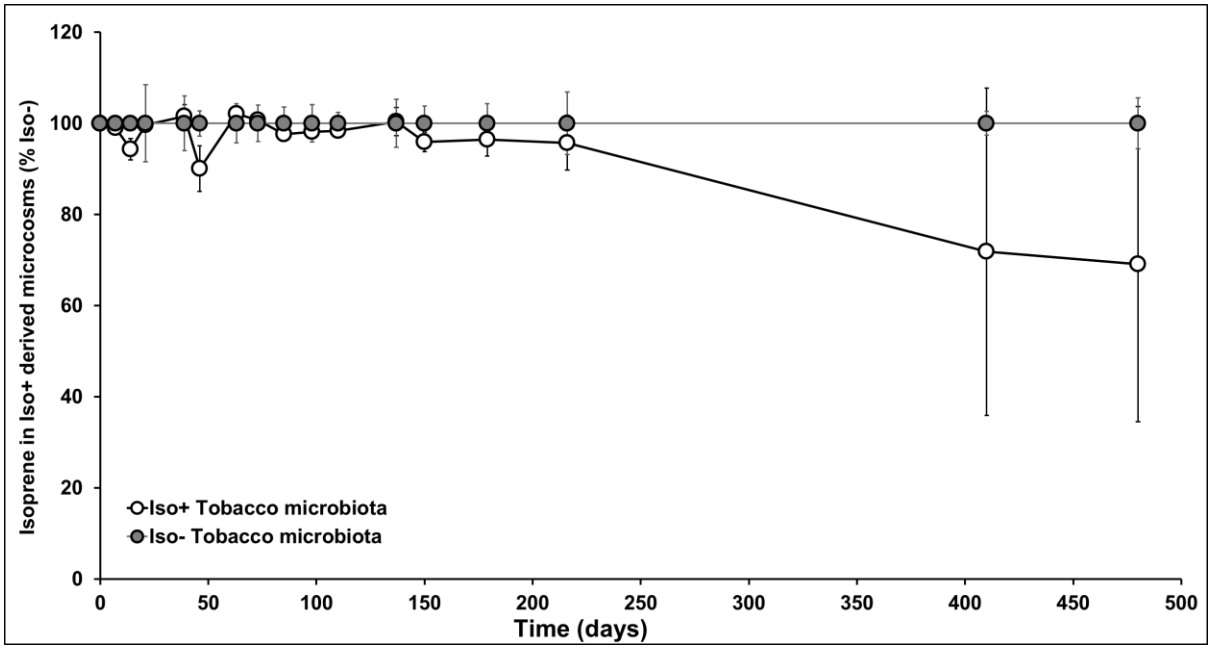


Figure 4.6 isoprene concentration over time for isoprene enrichments containing microbiota removed from isoprene-producing tobacco, compared to microcosms containing microbiota removed from isoprene-non-producing tobacco. Iso+ from isoprene-producing tobacco, Iso- from non-isoprene-producing tobacco; Error bars are standard error. n=3.

4.5 Discussion

Sequence analysis of leaf bacterial community data, for isoprene-producing and non-producing tobacco, revealed a lack of difference , and therefore any effect of isoprene on the leaf microbiota is not significant. The primary conclusion from this is that isoprene production is unlikely to be a major driving force for community composition on leaves. However, it is notable that some of the genera that increased, or appeared in the added isoprene experiment have been associated previously with isoprene production, including *Rhodococcus* and TM7, even though the mechanism by which TM7 is enriched on isoprene is largely unknown. This change, although not significant, suggests that it would be inappropriate to rule out any effect, only any significant effect, under these experimental conditions. That said, the appearance of *Rhodococcus* on the tobacco leaves is unsurprising, as *Rhodococcus* species have been shown previously to live on tobacco, where *R*. *fascians* can cause leafy gall disease (Goethals et al. 2001).

Foremost in determining the outcome, I would suggest that possibly (i) the hostile environment normally present on leaves (Lindow & Brandl 2003) and (ii) the nicotine which has antimicrobial properties (Pavia et al. 2000), present in the sticky exudate that finely coats tobacco leaves, would have a large effect. Although in the experiment the wind and rain, as well as some of the solar radiation effects are removed by the greenhouse environment, and therefore the environment is possibly less hostile than the natural systems. It is also worth remembering that due to the recent production, and the controlled use of the transgenic tobacco system (Vickers et al. 2011); firstly, isoprene utilising bacteria may not have had time to adapt to the tobacco environment, and

secondly, any bacteria which may have had the capacity to consume isoprene and live on tobacco leaves may not have been able to colonise the leaves in the laboratory environment in which they were grown. Also, clouding any outcome is the tendency for leaves to have high bacterial community variability, and often distinct communities, in small spatial and time scales (Thompson et al. 1993). This leaf-to-leaf variation could explain why genera (e.g. *Rhodococcus*) with members known to degrade isoprene appeared on some replicates, yet were absent on others. One limitation of this study was that due to large amounts of chloroplast and mitrochondrial DNA from the tobacco there was low actual sequencing depth, and future work should be done with either a more gentle leaf wash protocol, or preferably a much higher sequencing depth.

Considering the results of the tobacco community analysis, the lack of isoprene degradation in the tobacco leaf wash microcosms is hardly surprising, with no degradation over a 200 (probably 450) day period. Although the artificial (i.e. 10 ml of liquid is quite different to a leaf) environment, including the media choice, could have affected the community. Although the genera found on the tobacco also include species found to degrade isoprene under the same conditions from soil (Chapter 2), so you could expect some of them would be able to grow under the conditions in this experiment. The lack of degradation, let alone faster degradation, in the enrichments from the isoprene-producing tobacco bacterial communities further evidences that isoprene is not a major driver of community structure. This is supported by the isoprene-degradation enrichments from Willow, Poplar and Birch leaves, where the trees were a long term system with natural inoculation, and still shows a high variation in degradation rate between microcosms and a lack of degradation in many, suggesting that although

present on isoprene-producing tree leaves, isoprene-degrading bacteria are not uniform in quantity or potentially presence.

The decrease of isoprene in the tobacco leaf wash could have been due to a leak given to the length of time. However, due to what we know from the stochastic nature of the isoprene degradation of the tree leaf washings, it is possible it is the early signs of isoprene degradation, starting slowely due to the low biomass of the inoculum.

The inability to measure any degradation of isoprene in chambers containing leaves was most likely due to the experimental difficulty in trying to monitor degradation in the background of production. Although it did demonstrate that even when leaves are at a stable temperature, in the dark (i.e. when isoprene production is expected to be down regulated to near zero (Alves et al., 2014) and in this case was less than 1 pmol (isoprene) m^{-2} (leaf) min^{-1}), the isoprene production from the leaves still overwhelms any possible isoprene degradation by bacteria, supporting the idea that if isoprene degradation is happening, it is not significant.

Considering the possibility that isoprene degradation is not an important process on leaves, it is hardly surprising that the experiment to see if net production of isoprene from leaf systems could be decreased with isoprene-degrading bacteria failed to show any difference between treated and non-treated leaves. And additionally not suprising if you consider that the isolates the leaves were inoculated with may not have had the other physiological adaptions required for leaf surface life, for example: pigmented bacteria are more common on leaves due to solar radiation pressures (Lindow & Brandl

2003), and some leaf bacteria have developed specific leaf attachment mechanisms (Sequeira et al. 1977), and therefore it is unlikely, or at least unclear whether the isoprene degraders added to the leaves were either active or present at the time of sampling. One possible explanation for the limited isoprene degradation (and isoprene degraders) on leaves could be due to the relative speed of isoprene movement through soil, where the isoprene must follow a tortuous path, compared to leaves, where it is directly released into air; giving bacteria little chance to capture it.

4.6 Conclusion

In summary, although it is clear that bacteria capable of degrading isoprene exist on isoprene-producing leaves (and can be isolated, such as the 27 leaf isolates identified in Chapter 3), it may be unlikely that (i) isoprene is used as a major carbon source or that (ii) any significant proportion of isoprene is degraded on the leaves, and it is probable that the leaf community structure is likely driven by other forces.

Acknowledgements

Many thanks to Claudia Vickers for providing the transgenic tobacco.

Final discussion

This project set out to investigate isoprene degradation in the terrestrial environment, specifically to: i) determine the effect that soil and leaf dwelling bacteria have on the consumption of isoprene, specifically at atmospherically relevant concentrations; ii) investigate and identify the isoprene-degrading bacteria in soils and on leaves; iii) determine the effect of isoprene on the leaf and soil bacterial communities; and iv) investigate isoprene-degrading isolates in depth.

By measuring the changes in isoprene concentration in chambers spiked with15 ppb isoprene deployed in situ, the isoprene degradation rate of soils was shown to be 4.77 nmol m^{-2} h^{-1} in this study, leading to an adjustment of the global estimate of isoprene uptake of soils from 20.4 to just 0.03 Tg C yr^{-1} . This is in contrast to some of the previous literature, including Cleveland and Yavitt (1997); however it supports the figures from Gray et al. (2014) of 2.0 nmol m^{-2} h^{-1}. This will have some impact on climate models which had been updated to include effects of the soil sink on isoprene concentrations and therefore global temperature, which would now need reducing appropriately (Cleveland & Yavitt, 1998).

Through seasonal testing, it was discovered that there was no significant differences in the rate of isoprene degradation by soils. An unusual aspect as isoprene is that it is largely produced by the leaves of trees, especially deciduous trees, thus leading to seasonally different levels of isoprene (Guenther etl al., 2006). As you would expect that there would be seasonal differences in the rate of isoprene degradation if isoprene-degrading bacteria

were enriched during the warmer months where broadleaf trees (the largest isoprene source) have leaves (the site of isoprene emission) and are subject to higher temperatures (which drive isoprene emission), the lack of a seasonal effect suggests that the ability of soils to degrade isoprene is not driven by atmospheric isoprene.

Further, experiments in this study to investigate the effect of proximity to isoprene-producing trees on soil degradative capacity, using high levels of isoprene in serial enrichments support this conclusion. Firstly there was no significant difference in the isoprene degrading capacity between the location types, and secondly the soils displayed rapid degradation at high concentrations of isoprene (80% of 10^5 ppb in 80 hours), much faster than in situ degradation rates. This adds to the evidence that the isoprene-degrading bacteria in these soil communities are not relying on isoprene derived from the atmosphere. Taken together, these findings suggest that much of the isoprene in the soil must be coming from a non-atmospheric source, the most likely of which would be isoprene production within the soil. As far as we know trees do not naturally produce isoprene from their roots (Cinege et al., 2009), and we know that bacteria in soils are capable of producing isoprene (Murphy, 2011), suggesting a microbial isoprene cycle within the soil. This potential in soil isoprene cycle, and the ability of soils to rapidly degrade isoprene at concentrations that are higher than those found in the atmosphere would also go some way to explain why, in experiments where isoprene degradation rates of soils were determined using above naturally occurring atmospheric isoprene levels, the isoprene degradation rates were much higher than those preformed at environmentally relevant concentrations. It is important to note the limitations

of this study. For example, the isoprene degradation rates reported here were in a small geographic area (although supported by the others' work e.g. Gray et al. 2014), and therefore the adjustment of the global soil sink for isoprene is subject to those limitations, and is an initial estimate rather than an absolute value. Further valuable work on this topic would be to perform in situ degradation experiments at environmentally relevant concentrations of isoprene, or with no isoprene (as with Gray et al., 2014). With the community analyses, future work could concentrate on low concentration isoprene enrichments, in a similar style to this project in different soil types, and possibly with SIP. In this thesis cyclohexamide was used to prevent fungal growth in the isoprene enrichments, however as fungal biomas can be 1.5 to 4.3 fold bacterial biomass in soils (Ananyeva et al., 2006), it is could be usefull to include these in future studies (even though fungi show little response to isoprene enrichment (Gray et al., 2015)).

The various studies in this thesis demonstrated that the isoprene-degrading bacteria enriched depend on the isoprene concentration. At high isoprene concentrations genera such as *Rhodococcus* dominate, yet a low concentrations other bacteria like *Methylobacterium* are enriched. This points towards a two tiered system of high and low affinity isoprene degraders in the environment, or more probably a spectrum of affinities. Isoprene may be scavenged by high affinity microbes as a trace gas, sustaining bacteria when nutrients are limited, like with CO_2, H_2 and CH_4 (Greening et al., 2015; Quiza et al., 2014; Dunfield et al., 1999). These experiments also raised the question of the role of the deeply branching phylum, TM7, in isoprene consumption. TM7 was enriched to high levels in most soils when isoprene was added at high

concentrations and was found in higher abundance (although not statistically significant) on isoprene-producing tobacco compared with non-producing tobacco, yet no isoprene uptake into the DNA was found in the SIP experiments. It is possible that TM7 is using isoprene as an energy source (i.e. like *Gluconobacter oxydans* does with D-lactate in Sheng et al. 2015), rather than a carbon source, instead using other trace compounds as carbon sources.

Additionally, from the serial enrichment experiments it became evident that the bacterial communities derived from soils from different locations became more similar after isoprene enrichment, demonstrating that the community of isoprene-degrading bacteria in any location are somewhat similar. This observation is consistent with the analysis of isoprene-degrading bacteria from different locations; where, for the bacteria derived from this thesis, there were no clear splits in phylogeny, or *isoA* similarity based on location type suggesting a replicability of isoprene degrading communities upon enrichment.

Through indepth investigation of bacteria it was shown that many isoprene-degrading bacteria had a wide range of metabolites they were capable of using as sole carbon source. This lack of dependence on isoprene as a carbon source may suggest that isoprene is used mainly as an additional carbon source where available, and the lack of replicability of the genomic structure of the isoprene degrading operon and the lack of phylogenetic relationship between the *isoA* gene (important in isoprene degradation) and the *16S rRNA* gene (phylogenetically linked) suggests that isoprene-degrading genes may be subject to lateral transfer. This project lead to the isolation of a number of isoprene-degrading strains, and future work could expand the genomic analysis to more than the few selected in this work, and potentially the potentially

interesting *Stenotrophomonas* and *Bosea* strains (unfortunately the time limitations of this work meant that these could not be done in this project). These experiments also began to answer the TM7 question, showing a potential metabolic interaction with *Aeromicrobium*, although this was based on low depth metagenomics with one sample, greater depth and statistical power would be important for detailing the TM7 story.

Isoprene degradation by bacteria on leaves was studied due to leaves being the major site of isoprene release. The investigation of isoprene degradation by bacteria on, or from leaves demonstrated isoprene degradation by leaf bacteria, however the community analysis of the bacteria on isoprene-producing and non-producing tobbaco had limited results, with little statistical significance. Where bacteria were enriched from leaves, enrichment happened sporadically and was not correlated to the isoprene production status of the source leaves, therefore if there are differences in communities of isoprene degraders enrichable by these methods, isoprene is unlikely to be the driving force. The metagenetic analysis of the *16S rDNA* amplicons gave no significant differences, although this may have been an issue with sequencing depth. They did however demonstrate higher concentrations of TM7 on some of the isoprene-producing tobacco leaves, which may indicate an interaction there. This also reiterates the question of how TM7 is interacting with isoprene.

One clear output of this study was that if isoprene degraders are degrading isoprene on leaves, it is overwhelmed by the little isoprene the leaves produce in the dark, and therefore it is negligable in comparison to daytime rates of isoprene production, suggesting that there is little benefit in accurate quantification.

With the data from this thesis, a picture emerges that isoprene degradation on leaves is limited, with few isoprene-degrading bacteria on leaves, and the bacteria on leaves degrading little of the isoprene produced, with poor replicability between environments, and an inability to meaningfully support introduced isoprene-degraders. However this is a very tenuous argument due to having no statistical significance associated with it (of sameness), and therefore could be indicative only. Future work in this area should focus on using a non-tobacco model organism (as tobacco sap has antimicrobial properties (Pavia el al., 2009)), perhaps the aribadopsis from Loivamaki et al. 2007 or better still the poplar with silenced isoprene synthase like in Müller et al. 2009 alongside much deeper sequencing (due to chloroplast and mitrochondria contamination), and enrichments of leaf microbiota should have larger innoculums, and measurements of isoprene degraders introduced to plants could be accompanied by measures of survival.

References

Acuña Alvarez, L. Exton, D. A. Timmis, K. N. Suggett, D. J. and McGenity, T. J. (2009). Characterization of marine isoprene-degrading communities. *Environmental Microbiology*, **11**, 3280–91.

Altschul, S.F. Gish, W. Miller, W. Myers, E. W. and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403-410.

Alves, E. G. Harley, P. Gonçalves, J. F. de C. Moura, C. E. da S. and Jardine, K. (2014). Effects of light and temperature on isoprene emission at different leaf developmental stages of *Eschweilera coriacea* in central Amazon. *Acta Amazonica*, **44**, 9–18.

Ananyeva, N. D. Susyan, E. A. Chernova, O. V. Chernov, I. Y. and Makarova, O. L. (2006). The ratio of fungi and bacteria in the biomass of different types of soil determined by selective inhibition. *Microbiology*, **75**, 702–707.

Armstrong, A. (2013). Atmospheric chemistry: Isoprene and agriculture. *Nature Geoscience*, **6**, 513–513.

Arneth, A. Miller, P. A. Scholze, M. Hickler, T. Schurgers, G. Smith, B. and Prentice, I. C. (2007). CO_2 inhibition of global terrestrial isoprene emissions: Potential implications for atmospheric chemistry. *Geophysical Research Letters*, **34**, 18813.

Arneth, A. Schurgers, G. Lathiere, J. Duhl, T. Beerling, D. J. Hewitt, C. N. Martin, M. Guenther, A. (2011). Global terrestrial isoprene emission

models: sensitivity to variability in climate and vegetation. *Atmospheric Chemistry and Physics*, **11**, 8037–8052.

Ashworth, K. Wild, O. and Hewitt, C. N. (2013). Impacts of biofuel cultivation on mortality and crop yields. *Nature Climate Change,* **3**, 492–496.

Aslam, S. (2012), personal communication “DNA extraction from Leaf microbiota” *University of Essex*

Atkinson, R. (1990). Gas-phase tropospheric chemistry of organic compounds: A review. *Atmospheric Environment,* **24**, 1–41.

Atkinson, R. and Arey, J. (2003.) Atmospheric degradation of volatile organic compounds. *Chemical Reviews*, **103**, 4605–38.

Aziz, R. K. Bartels, D. Best, A. A. DeJongh, M. Disz, T. Edwards, R. A. Formsma, K. Gerdes, S. Glass, E. M. Kubal, M. Meyer, F. Olsen, G. J. Olson, R. Osterman, A. L. Overbeek, R. A. McNeil, L. K. Paarmann, D. Paczian, T. Parrello, B. Pusch, G. D. Reich, C. Stevens, R. Vassieva, O. Vonstein, V. Wilke, A. and Zagnitko, O. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, **9**, 75.

Bäcka, J. Aaltonena, H. Hellénc, H. Kajosb, M. K. Patokoskib, J. Taipaleb, R. Pumpanena, J. and Heinonsalod, J. (2015). Variable emissions of microbial volatile organic compounds (MVOCs) from root-associated fungi isolated from Scots pine. *Atmospheric Environment,* **44**, 3651-3659

Baldocchi, D. Guenther, A. Harley, P. Klinger, L. Zimmerman, P. Lamb, B. and Westberg, H. (1995). The fluxes and air chemistry of isoprene above a deciduous hardwood forest. *Philosophical Transactions: Physical Sciences and Engineering*, **351**, 269–296.

Banerjee, A. and Sharkey, T. D. (2014). Methylerythritol 4-phosphate MEP pathway metabolic regulation. *Natural Product Reports*, **31**, 1043–55.

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. and Pevzner, P. A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, **19**, 455–477.

Beattie, G. A. and Lindow, S. E. (1999). Bacterial colonization of leaves: a spectrum of strategies. *Phytopathology*, **89**, 353–359.

Behnke, K. Ehlting, B. Teuber, M. Bauerfeind, M. Louis, S. Hänsch, R. Polle, A. Bohlmann, J. and Schnitzler, J.P. (2007). Transgenic, non-isoprene emitting poplars don't like it hot. *The Plant Journal : For Cell and Molecular Biology*, **51**, 485–99.

Behnke, K. Grote, R. Brüggemann, N. Zimmer, I. Zhou, G. Elobeid, M. Janz, D. Polle, A. and Schnitzler, J.-P. (2012). Isoprene emission-free poplars--a chance to reduce the impact from poplar plantations on the atmosphere. *The New Phytologist*, **194**, 70–82.

Belfiore C, Castellano P. and Vignolo G (2007) Reduction of *Escherichia coli* population following treatment with bacteriocins from lactic acid bacteria and chelators. *Food Microbiology.* **24**, 223-229.

Bringel, F. and Couée, I. (2015) Pivotal roles of phyllosphere microorganisms at the interface between plant functioning and atmospheric trace gas dynamics. *Frontiers in Microbiology*. **6**, 486

Brettin, T. Davis, J. J. Disz, T. Edwards, R. A. Gerdes, S. Olsen, G. J. Olson, R. Overbeek, R. Parrello, B. Pusch, G. D. Shukla, M. Thomason, J. A. Stevens, R. Vonstein, V. Wattam, A. R. and Xia, F. (2015). RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports*, **5**, 8365.

Brown, C. T. Hug, L. A. Thomas, B. C. Sharon, I. Castelle, C. J. Singh, A. Wilkins, M. J. Wrighton, K. C. Williams, K. H. and Banfield, J. F. (2015). Unusual biology across a group comprising more than 15% of domain Bacteria*.* *Nature*, **523**, 208–11.

Bull, I. D. Parekh, N. R. Hall, G. H. Ineson, P. and Evershed, R. P. (2000). Detection and classification of atmospheric methane oxidizing bacteria in soil. *Nature*, **405**, 175–8.

Cinege, G. Louis, S. Hänsch, R. and Schnitzler, J.-P. (2009). Regulation of isoprene synthase promoter by environmental and internal factors. *Plant Molecular Biology*, **69**, 593–604.

Claeys, M. Graham, B. Vas, G. Wang, W. Vermeylen, R. Pashynska, V. Cafmeyer, J. Guyon, P. Andreae, M. O. Artaxo, P. and Maenhaut, W. (2004). Formation of secondary organic aerosols through photooxidation of isoprene. *Science*, **303**, 1173–6.

Cleveland, C. C. and Yavitt, J. B. (1997). Consumption of atmospheric isoprene in soil. *Geophysical Research Letters*, **24**, 2379-2382

Cleveland, C. C. and Yavitt, J. B. (1998). Microbial consumption of atmospheric isoprene in a temperate forest soil. *Applied and Environmental Microbiology*, **64**, 172–177.

Collins, W. J. Derwent, R. G. Johnson, C. E. and Stevenson, D. S. (2002). The Oxidation of Organic Compounds in the Troposphere and Their Global Warming Potentials. *Climatic Change*, **52**, 453–479.

Conrad, R. (2009). The global methane cycle: recent advances in understanding the microbial processes involved. *Environmental Microbiology Reports*, **1**, 285–292.

Cools N, De Vos B, (2010). Sampling and Analysis of Soil. Manual Part X. *Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests*, UNECE, ICP Forests, Hamburg.

Cordoba, E. Salmi, M. and León, P. 2009. Unravelling the regulatory mechanisms that modulate the MEP pathway in higher plants. *Journal of Experimental Botany*, **60**, 2933–43.

Crombie, A. T. El Khawand, M. Rhodius, V. A. Fengler, K. A. Miller, M. C. Whited, G. M. McGenity, T. J. and Murrell, J. C. (2015). Regulation of plasmid-encoded isoprene metabolism in *Rhodococcus* , a representative of an important link in the global isoprene cycle. *Environmental Microbiology*, **17**, 3314-3329.

Curtis, T. P. Sloan, W. T. and Scannell, J. W. (2002). Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 10494–9.

Dani, K. G. S. Jamie, I. M. Prentice, I. C. and Atwell, B. J. (2014). Increased Ratio of Electron Transport to Net Assimilation Rate Supports Elevated Isoprenoid Emission Rate in Eucalypts under Drought. *Plant Physiology*, **166**, 1059–1072.

Donoso, L. Romero, R. Rondón, A. Fernandez, E. Oyola, P. and Sanhueza, E. (1996). Natural and anthropogenic C2 to C6 hydrocarbons in the central-eastern Venezuelan atmosphere during the rainy season. *Journal of atmospheric chemistr*y, **25**, 201-214.

Duane, M. Poma, B. Rembges, D. Astorga, C. and Larsen, B. (2002). Isoprene and its degradation products as strong ozone precursors in Insubria, Northern Italy. *Atmospheric Environment*, **36**, 3867–3879.

Dunfield, P. F. Liesack, W. Henckel, T. Knowles, R. and Conrad, R. (1999). High-affinity methane oxidation by a soil enrichment culture containing a type II *Methanotroph*. *Applied and Environmental Microbiology*, **65**, 1009–14

Dyas, A. Boughton, B. J. and Das, B. C. (1983). Ozone killing action against bacterial and fungal species ; microbiological testing of a domestic ozone generator. *Journal of Clinical Pathology*, **36**, 1102–1104.

El Khawand, M. (2014). Bacterial Degradation of Isoprene in the Terrestrial Environment. *Doctoral Thesis, University of East Anglia, UK.*

El Khawand, M. Crombie, A.T. Johnston, A. Vavlline, D. V. McAuliffe, J. C. Latone, J. A. Primak, Y. A. Lee, S. K. Whited, G. M. McGenity, T. J. and Murrell, J. C. (2016) Isolation of isoprene degrading bacteria from soils, development of isoA gene probes and identification of the active isoprene-

degrading soil community using DNA-stable isotope probing. *Environmental Microbiology,* **10**, 1462-2920

Ensign, S. A. Hyman, M. R. and Arp, D. J. (1992). Cometabolic degradation of chlorinated alkenes by alkene monooxygenase in a propylene-grown *Xanthobacter* strain. *Applied and Environmental Microbiology*, **58**, 3038–46.

Ewers, J. Freier-Schröder, D. and Knackmuss, H.-J. (1990). Selection of trichloroethene TCE degrading bacteria that resist inactivation by TCE. *Archives of Microbiology*, **154**, 410–413.

Fahy, A. McGenity, T. J. Timmis, K. N. and Ball A. S. (2006) Heterogeneous aerobic benzene-degrading communities in oxygen-depleted groundwaters. *FEMS Microbial Ecology*, **58**, 260-270.

Fall, R. and Copley, S. D. (2000). Bacterial sources and sinks of isoprene, a reactive atmospheric hydrocarbon. *Environmental Microbiology*, **2**, 123–130.

Fall, R. and Monson, R. K. (1992). Isoprene emission rate and intercellular isoprene concentration as influenced by stomatal distribution and conductance. *Plant Physiology*, **100**, 987–92.

Fan, J. and Zhang, R. (2004). Atmospheric Oxidation Mechanism of Isoprene. *Environmental Chem*istry, **1**, 140.

Fehsenfeld, F. Calvert, J. Fall, R. Goldan, P. Guenther, A. B. Hewitt, C. N. Lamb, B. Liu, S. Trainer, M. Westberg, H. and Zimmerman, P. 1992. Emissions of volatile organic compounds from vegetation and the

implications for atmospheric chemistry. *Global Biogeochemical Cycles*, **6**, 389–430.

Felsenstein J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution,* **39**, 783-791.

Ferrari BC, Binnerup SJ and Gillings M (2005) Microcolony Cultivation on a Soil Substrate Membrane System Selects for Previously Uncultured Soil Bacteria. *Applied and Environmental Microbiology*. **71**, 8714-8720.

Field, D. Tiwari, B. Booth, T. Houten, S. Swan, D. Bertrand, N. and Thurston, M. (2006). Open software for biologists: from famine to feast. *Nature Biotechnology*, **24**, 801–803.

Geron, C. Guenther, A. Sharkey, T. and Arnts, R. R. (2012). Temporal variability in basal isoprene emission factor. *Tree Physiology*, **12**, 799-805.

Geng, F. Tie, X. Guenther, A. Li, G. Cao, J. and Harley, P. (2011). Effect of isoprene emissions from major forests on ozone formation in the city of Shanghai, China. *Atmospheric Chemistry and Physics*, **11**, 10449-10459.

Giebler, J. Y. Wick, L. Y. Chatzinotas, A. Harms H. (2013). Alkane-degrading bacteria at the soil–litter interface: comparing isolates with T-RFLP-based community profiles. *FEMS Microbial Ecology*, **86**, 45-58

Goldan, P. D. Kuster, W. C. Fehsenfeld, F. C. and Montzka, S. A. (1995). Hydrocarbon measurements in the southeastern United States: The Rural Oxidants in the Southern Environment ROSE Program (1990). *Journal of Geophysical Research: Atmospheres*, **100**, 25945 – 25963.

Goldstein, H. Goulden, M. L. Munger, J. W. Wofsy, C. and Geron, D. (1998). Seasonal course of isoprene emissions from a midlatitude deciduous forest. *Journal of Geophysical Research*, **103**, 31045 – 31056.

Gounaris, K. Brain, A. R. R. Quinn, P. J. and Williams, W. P. (1984). Structural reorganisation of chloroplast thylakoid membranes in response to heat-stress. Biochimica et Biophysica Acta BBA - *Bioenergetics*, **766**, 198–208.

Granier, C. Petron, G. Muller, J. F. and Brasseur G. (2000). The impact of natural and anthropogenic hydrocarbons on the tropospheric budget of carbon monoxide. *Atmospheric Environment,* **34**, 5255–5270.

Gray, C. M. Helmig, D. and Fierer, N. (2015). Bacteria and fungi associated with isoprene consumption in soil. *Elementa: Science of the Anthropocene*, **3**, 53.

Gray, C. M. Monson R.K. and Fierer N. (2014). Biotic and abiotic controls on biogenic volatile organic compound fluxes from a subalpine forest floor. *Journal of Geophysical Research: Biogeosciences*, **119**, 547–556.

Greening, C. Carere, C. R. Rushton-Green, R. Harold, L. K. Hards, K. Taylor, M. C. Morales, S. E. Stott, M. B. and Cook, G. M. (2015). Persistence of the dominant soil phylum *Acidobacteria* by trace gas scavenging. *Proceedings of the National Academy of Sciences*, **112**, 10497–10502.

Griffiths, R. I. Whiteley, A. S. O'Donnell, A. G. and Bailey, M. J. (2000). Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Applied and Environmental Microbiology*, **66**, 5488–5491.

Guenther, A. Hewitt, C. N. Erickson, D. Fall, R. Geron, C. Graedel, T. Harley, P. Klinger, L. Lerdau, M. Mckay, W. A. Pierce, T. Scholes, B. Steinbrecher, R. Tallamraju, R. Taylor, J. and Zimmerman, P. (1995). A global model of natural volatile organic compound emissions. *Journal of Geophysical Research*, **100**, 8873.

Guenther, A. Hewitt, C. N. Erickson, D. Fall, R. Geron, C. Graedel, T. Harley, P. Klinger, L. Lerdau, M. Mckay, W. A. Pierce, T. Scholes, B. Steinbrecher, R. Tallamraju, R. Taylor, J. and Zimmerman, P. (1995), A global model of natural volatile organic compound emissions *Journal of Geophysical Research: Atmospheres*, **100**, 8873-8892.

Guenther, A. Karl, T. Harley, P. Wiedinmyer, C. Palmer, P. I. and Geron, C. (2006). Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmospheric Chemistry and Physics Discussions*, **6**, 107–173.

Haahr M. School of Computer Science and Statistics at Trinity College, Dublin, Ireland (2008). True random sequence generation. URL http://www.random.org/sequences/. Last retrieved 23.11.15

Han, J. I. Spain, J. C. Leadbetter, J.R. Ovchinnikova, G. Goodwin, L.A. Han C. S., Woyke, T. Karen, W. Davenport, K. W. (2013). Genome of the Root-Associated Plant Growth-Promoting Bacterium *Variovorax paradoxus* Strain EPS, *Genome Announcements*, **1**, 4–5.

Hannon Lab. FASTX Toolkit. URL: http://hannonlab.cshl.edu/fastx_toolkit/index.html. Last retrieved 23.11.15

Hanson, D. T. Swanson, S. Graham, L. E. and Sharkey, T. D. (1999). Evolutionary significance of isoprene emission from mosses, temperatures frequently experienced by leaves at the top. *American Journal of Botany*, **86**, 634–639.

Hardacre, C. J. Palmer, P. I. Baumanns, K. Rounsevell, M. and Murray-Rust, D. (2013). Probabilistic estimation of future emissions of isoprene and surface oxidant chemistry associated with land-use change in response to growing food needs. *Atmospheric Chemistry and Physics*, **13**, 5451–5472.

Harley, P. C. Monson, R. K. and Lerdau, M. T. (1999). Ecological and evolutionary aspects of isoprene emission from plants. *Oecologia*, **118**, 109–123.

Harley, P. Guenther, A. and Zimmerman, P. (1996). Effects of light, temperature and canopy position on net photosynthesis and isoprene emission from sweetgum (*Liquidambar styraciflua*) leaves. *Tree Physiology*, **16**, 25–32.

He, X. McLean, J. S. Edlund, A. Yooseph, S. Hall, A. P. Liu, S.-Y. Dorrestein, P. C. Esquenazi, E. Hunter, R. C. Cheng, G. Nelson, K. E. Lux, R. and Shi, W. (2015). Cultivation of a human-associated TM7 phylotype reveals a reduced genome and epibiotic parasitic lifestyle. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 244–9.

Heald, C. L. Wilkinson, M. J. Monson, R. K. Alo, C. A. Wang, G. and Guenther, A. (2009). Response of isoprene emission to ambient CO_2 changes and implications for global budgets. *Global Change Biology*, **15**, 1127–1140.

Hernandez-Eugenio, G. Fardeau, M. L. Garcia, J.L. and Ollivier, B. (2015). Bergey's Manual of Systematics of Archaea and Bacteria. (W. B. Whitman, F. Rainey, P. Kämpfer, M. Trujillo, J. Chun, P. DeVos, B. Hedlund, & S. Dedysh, Eds.) (Vol. 1–6). Genus III *Aeromicrobium*, pp. 1258-1267. *Chichester, UK: John Wiley & Sons, Ltd*.

Hewitt, C. N. MacKenzie, A. R. Di Carlo, P. Di Marco, C. F. Dorsey, J. R. Evans, M. Fowler, D. Gallagher, M. W. Hopkins, J. R. Jones, C. E. Langford, B. Lee, J. D. Lewis, a C. Lim, S. F. McQuaid, J. Misztal, P. Moller, S. J. Monks, P. S. Nemitz, E. Oram, D. E. Owen, S. M. Phillips, G. J. Pugh, T. A. M. Pyle, J. A. Reeves, C. E. Ryder, J. Siong, J. Skiba, U. and Stewart, D. J. (2009). Nitrogen management is essential to prevent tropical oil palm plantations from causing ground-level ozone pollution. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 18447–51.

Hewitt, C. N. (n.d.). Database on volatile organic compound emissions. *Lancaster University*.

Hibbing, M. E. Fuqua, C. Parsek, M. R. and Peterson, S. B. (2010). Bacterial competition: surviving and thriving in the microbial jungle. *Nature Reviews Microbiology*, **8**, 15–25.

Hoagland, D. R. and Arnon, D. I. (1938). Water-culture method for growing plants without soil. *California Agricultural Experiment Station Circular*, **347**, 1884-1949.

Hou, C. T. Patel, R. N. Laskin, N. Barnabe, N. and Barist, I. (1981). Epoxidation and hydroxylation of C4 - and C5 -branched-chain alkenes and alkanes by *Methanotrophs*. *Industrial Microbiology*, **23**, 477–482.

Huang, X. and Madan, A. (1999). CAP3: A DNA sequence assembly program. *Genome Research*, **9**, 868–77.

Hugenholtz P, Tyson GW, Webb RI, Wagner AM and Blackall LL (2001) Investigation of candidate division TM7, a recently recognized major lineage of the domain Bacteria with no known pure-culture representatives. *Applied and Environmental Microbiology.* **67**, 411–419.

Huixiang, W. Kiang, C. S. Xiaoyan, T. Xiuji, Z. and Chameides, W. L. (2005). Surface ozone : A likely threat to crops in Yangtze delta of China. *Atmospheric Environment,* **39**, 3843–3850.

Ikeda, S. Kaneko, T. Okubo, T. Rallos, L. E. Eda, S. Mitsui, H. Sato, S. Nakamura, Y. Tabata, S. and Minamisawa, K. (2009). Development of a bacterial cell enrichment method and its application to the community analysis in soybean stems. *Microbial Ecology*, **58**, 703-714.

Illumina Inc. (2013). 16S Metagenomic Sequencing Library Preparation - Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System. *16S Metagenomic Sequencing Library Preparation Manual*, 1–28.

IPCC United Nations Environment Programme. (2001). Trace Gases: Current Observations, Trends and Budgets. *Third Assessment Report - Climate Change*.

Jacobs, J. L. and Sundin, G. W. (2001), Effect of Solar UV-B Radiation on a Phyllosphere Bacterial Community. *Applied and Environmental Microbiology*, **67**, 5488-5496

Jenkin, M. E. and Clemitshaw, K. C. (2002). Ozone and other secondary photochemical pollutants: chemical processes governing their formation in the planetary boundary layer. *Developments in Environmental Science*, **1**, 285–338.

Johan, E. T. Janssen, D. B. Utilization, I. and Strain, S. P. (1998). Dichloroepoxyethane Is Involved in isoprene Utilization by *Rhodococcus* sp . Strain AD45, *Applied and Environmental Microbiology,* **64**, 2800–2805.

Johnston, A. (2014). Molecular Ecology of Marine isoprene Degradation. *Doctoral Thesis, University of East Anglia, UK.*

Joshi, N. A. and Fass, J. N. (2011). Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). URL https://github.com/najoshi/sickle. Last retrieved 23.11.15

Jukes T.H. and Cantor C.R. (1969). Evolution of protein molecules. (Munro HN, editor), Mammalian Protein Metabolism, *Academic Press, New York*.

Julsing, M. K. Rijpkema, M. Woerdenbag, H. J. Quax, W. J. and Kayser, O. (2007). Functional analysis of genes involved in the biosynthesis of isoprene in *Bacillus subtilis*. *Applied Microbiology and Biotechnology*, **75**, 1377–84.

Kanawade, V. P. Jobson, B. T. Guenther, A. B. Erupe, M. E. Pressley, S. N. Tripathi, S. N. and Lee, S.H. (2011). Isoprene suppression of new particle

formation in a mixed deciduous forest. *Atmospheric Chemistry and Physics*, **11**, 6013–6027.

Kanawade, V. P. Jobson, B. T. Guenther, a. B. Erupe, M. E. Pressley, S. N. Tripathi, S. N. and Lee, S.-H. (2011). Isoprene suppression of new particle formation in a mixed deciduous forest. *Atmospheric Chemistry and Physics*, **11**, 6013–6027.

Kanehisa, M. and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28, 27–30.

Kantor, RS. Wrighton K. C. Handley K. M. Sharon I. Hug L. A. Castelle C. J. Thomas B C. Benfield J. K. (2013). Small Genomes and Sparse Metabolisms of Sediment-Associated Bacteria from Four Candidate Phyla. *Molecular Biology*, **4**, 708-713.

Kesselmeier, J. and Staudt, M. (1999). Biogenic Volatile Organic Compounds (VOC): An Overview on Emission , Physiology and Ecology. *Journal of Atmospheric Chemistry*, **33**, 23–88.

Kesselmeier, J. Kuhn, U. Wolf, A. Andreae, M. O. Ciccioli, P. Brancaleoni, E. and de Oliva, T. (2000). Atmospheric volatile organic compounds (VOC) at a remote tropical forest site in central Amazonia*. Atmospheric Environment*, **34**, 4063-4072.

Kiendler-Scharr, A. Wildt, J. Dal Maso, M. Hohaus, T. Kleist, E. Mentel, T. F. Tillmann, R. Uerlings, R. Schurr, U. and Wahner, A. (2009). New particle formation in forests inhibited by isoprene emissions. *Nature*, **461**, 381–4.

King, G. M. (1992). Ecological aspects of methane oxidation, a key determinant of global methane dynamics. *Advances in Microbial Ecology*, **12**, 431–468.

Klindworth, A. Pruesse, E. Schweer, T. Peplies, J. Quast, C. Horn, and M. Glöckner, F. O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research*, **7**, 41

Kuzma, J. Nemecek-Marshall, M. Pollock, W. H. and Fall, R. (1995). Bacteria produce the volatile hydrocarbon isoprene. *Current Microbiology*, **30**, 97-103.

Kuzuyama, T. (2002). Mevalonate and nonmevalonate pathways for the biosynthesis of isoprene units. *Bioscience, Biotechnology, and Biochemistry,* **66**, 1619–27.

Laothawornkitkul, J. Paul, N. D. Vickers, C. E. Possell, M. Mullineaux, P. M. Hewitt, C. N. and Taylor, J. E. (2008). The role of isoprene in insect herbivory, *Plant Signaling & Behaviour*, **3**, 1141–1142.

Leahy, J. G. Batchelor, P. J. and Morcomb, S. M. (2003). Evolution of the soluble diiron monooxygenases. *FEMS Microbiology Reviews*, **27**, 449–479.

Lewis, A. C. Bartle, K. D. & Rattner, L. (1997). High-speed isothermal analysis of atmospheric isoprene and DMS using on-line two-dimensional gas chromatography. *Environmental science & technology*, **31**, 3209-3217.

Li, L. and Wang, X. (2012). Seasonal and diurnal variations of atmospheric non-methane hydrocarbons in Guangzhou, China. *International Journal of Environmental Research and Public Health*, **9**, 1859–73.

Lightfield, J. Fram, N. R. and Ely, B. (2011). Across Bacterial Phyla, Distantly-Related Genomes with Similar Genomic GC Content Have Similar Patterns of Amino Acid Usage. *PLoS ONE*, **6**, e17677.

Lindow, S. E. and Brandl, M. T. (2003). Microbiology of the Phyllosphere. *Applied and Environmental Microbiology*, **69**, 1875–1883.

Logan, B. A. Monson, R. K. and Potosnak, M. J. (2000). Biochemistry and physiology of foliar isoprene production. *Trends in Plant Science*, **5**, 477–481.

Loivamaki, M. Gilmer, F. Fischbach, R. J. Sorgel, C. Bachl, A. Walter, A. and Schnitzler J. P. (2007) *Arabidopsis*, a Model to Study Biological Functions of Isoprene Emission? *American Society of Plant Biologists*, **144**, 1066-1078

Loivamäki, M. Mumm, R. Dicke, M. and Schnitzler, J. P. (2008). Isoprene interferes with the attraction of bodyguards by herbaceous plants. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 17430–5.

Loreto, F. and Sharkey, T. D. (1993). Isoprene emission by plants is affected by transmissible wound signals. *Plant, Cell and Environment*, **16**, 563–570.

Loreto, F. Pinelli, P. Manes, F. and Kollist, H. (2004). Impact of ozone on monoterpene emissions and evidence for an isoprene-like antioxidant action of monoterpenes emitted by Quercus ilex leaves. *Tree Physiology*, **24**, 361–7.

Loreto, F. Velikova, V. Nazionale, C. Vegetali, E. and Km, V. S. (2001). Isoprene Produced by Leaves Protects the Photosynthetic Apparatus

against Ozone Damage , Quenches Ozone Products , and Reduces Lipid Peroxidation of Cellular Membranes. *Plant Physiology*, **127**, 1781–1787.

Luo C, Xie S, Sun W, Xiangdong L and Cupples AM (2009) Identification of a Novel Toluene-Degrading Bacterium from the Candidate Phylum TM7, as Determined by DNA Stable Isotope Probing. *Applied and Environmental Microbiology*. **75**, 4644-4647.

Mabberley, D. J. (1987). The Plant-Book: A Portable Dictionary of the Higher Plants: 9780521340601.

Madden M, Song C, Tse K and Wong A (2004) The Inhibitory Effect of EDTA and Mg2+ on the Activity of NADH Dehydrogenase in Lysozyme Lysis. *Journal of Experimental Microbiology and Immunology*. **5**, 8-15.

Mahé, F. Rognes, T. Quince, C. de Vargas, C. and Dunthorn, M. (2014). Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ*, **2**, e593.

Malaysian Palm Oil Board MPOB. (2014). http://www.mpob.gov.my/palm-info/environment/520-achievements retrieved 03 December 2015

Masella, A. P. Bartram, A. K. Truszkowski, J. M. Brown, D. G. and Neufeld, J. D. (2012). PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics*, **13**, 31.

Mayrhofer, S. Mikoviny, T. Waldhuber, Wagner, A. O. Innerebner, G. Franke-Whittle, I. H. D. Märk, T. D. Hansel, A. Insam, H. (2006). Microbial community related to volatile organic compound (VOC) emission in household biowaste. *Environmental Microbiology*, **8**, 1960-1974.

Meyer, F. Paarmann, D. D'Souza, M. Olson, R. Glass, E. M. Kubal, M. Paczian, T. Rodriguez, A. Stevens, R. Wilke, A. Wilkening, J. and Edwards, R. A. (2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*, **9**, 386.

Monson, R. K. and Fall, R. (1989). isoprene emission from aspen leaves : influence of environment and relation to photosynthesis and photorespiration. *Plant Physiology,* **90**, 267–274.

Monson, R. K. Harley, P. C. Litvak, M. E. Wildermuth, M. Guenther, A. B. Zimmerman, P. R. and Fall, R. (1994). Environmental and developmental controls over the seasonal pattern of isoprene emission from aspen leaves. *Oecologia*, **99**, 260–270.

Müller, A. Kaling, M. Faubert, P. Gort, G. Smid, H. M. Van Loon, J. J. Dicke, M. Kanawati, B. Schmitt-Kopplin, P. Polle, A. Schnitzler, J.-P. Rosenkranz, M. and Rosenkranz, M. (2015). Isoprene emission by poplar is not important for the feeding behaviour of poplar leaf beetles. *BMC Plant Biology*, **15**, 165.

Müller, J.-F. Peeters, J. and Stavrakou, T. (2014). Fast photolysis of carbonyl nitrates from isoprene. *Atmospheric Chemistry and Physics* **14**, 2497–2508.

Murphy, G. P. (2011). Investigating and increasing the production of isoprene in E. coli and Bacillus species, MSc Thesis, *University of Essex, UK*.

Nadalig, T. Ul Haque, M. F. Roselli, S. Schaller, H. Bringel, H. and Vuilleumier, S. (2011) Detection and isolation of chloromethane - degrading bacteria

from the *Arabidopsis thaliana* phyllosphere, and characterization of chloromethane utilization genes. *FEMS Microbial Ecology*. **77**, 438-448

National Center for Biotechnology Information. PubChem Compound Database; CID=6557, https://pubchem.ncbi.nlm.nih.gov/compound/6557 accessed Dec 2, 2015.

Olcese, L. E. Penner, J. E. and Sillman, S. (2007). Development of a secondary organic aerosol formation mechanism : comparison with smog chamber experiments and atmospheric measurements. *Atmospheric Chemistry and Physics Discussions*, **7**, 8361–8393.

Overbeek, R. Olson, R. Pusch, G. D. Olsen, G. J. Davis, J. J. Disz, T. Edwards, R. A. Gerdes, S. Parrello, B. Shukla, M. Vonstein, V. Wattam, A. R. Xia, F. and Stevens, R. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research*, **42**, 206–214.

Pacifico, F. Harrison, S. P. Jones, C. D. and Sitch, S. 2009. Isoprene emissions and climate. *Atmospheric Environment*, **43**, 6121–6135.

Pagani, I. Liolios, K. Jansson, J. Chen, I. M. A. Smirnova, T. Nosrat, B. Markowitz M. and Kyrpides, N. C. (2012). The Genomes OnLine Database (GOLD) v.4: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Research*, **40**, 571–579.

Parte, A. C. (2014). LPSN--list of prokaryotic names with standing in nomenclature. *Nucleic Acids Research*, **42**, D613–6.

Paulot, F. Crounse, J. D. Kjaergaard, H. G. Kürten, A. St Clair, J. M. Seinfeld, J. H. and Wennberg, P. O. (2009). Unexpected epoxide formation in the gas-phase photooxidation of isoprene. *Science*, **325**, 730-733.

Paulson, S. E. and Seinfeld, J. H. (1992). Development and evaluation of a photooxidation mechanism for isoprene. *Journal of Geophysical Research*, **97**, 20703.

Pavia, C. S. Pierre, A. and Nowakowski, J. (2000). Antimicrobial activity of nicotine against a spectrum of bacterial and fungal pathogens. *Journal of Medical Microbiology*, **49**, 675–6.

Pegoraro, E. (2004). Environmental Control of Isoprene Emission : from Leaf to Canopy Scale. PhD Thesis, *University of Edinburgh.*

Pegoraro, E. Abrell, L. Van Haren, J. Barron-Gafford, G. Grieve, K. A. Malhi, Y. Murthy, R. and Lin, G. (2005). The effect of elevated atmospheric CO_2 and drought on sources and sinks of isoprene in a temperate and tropical rainforest mesocosm. *Global Change Biology*, **11**, 1234–1246.

Pfister, G. G. Emmons, L. K. Hess, P. G. Lamarque, J. F. Orlando, J. J. Walters, S. Guenther, A. Palmer, P. I. and Lawrence, P. J. (2008). Contribution of isoprene to chemical budgets: A model tracer study with the NCAR CTM MOZART-4. *Journal of Geophysical Research*, **113**, 1–21.

Pulido, P. Perello, C. and Rodriguez-Concepcion, M. (2012). New insights into plant isoprenoid metabolism. *Molecular Plant*, **5**, 964–7.

Purves, D. W. Caspersen, J. P. Moorcroft, P. R. Hurtt, G. C. and Pacala, S. W. (2004). Human-induced changes in US biogenic volatile organic compound

emissions: evidence from long-term forest inventory data. *Global Change Biology,* **10**, 1737–1755.

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

R Development Core Team (2008). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.*

Ramirez, K. S. Lauber, C. L. and Fierer, N. (2009). Microbial consumption and production of volatile organic compounds at the soil-litter interface. *Biogeochemistry,* **1**, 97-107

Rappé, M. S. and Giovannoni, S. J. (2003). The uncultured microbial majority. *Annual Review of Microbiology*, **57**, 369–94.

Rasmussen, R. A. and Went, F. W. (1965). Volatile organic material of plant origin in the atmosphere. *Proceedings of the National Academy of Sciences of the United States of America*, **53**, 215–20.

Rasulov, B. Huve, K. Bichele, I. Laisk, A. and Niinemets, U. (2010). Temperature Response of Isoprene Emission in Vivo Reflects a Combined Effect of Substrate Limitations and Isoprene Synthase Activity: A Kinetic Analysis. *Plant Physiology*, **154**, 1558–1570.

Relhum. http://hyperphysics.phy-astr.gsu.edu/hbase/kinetic/relhum.html. (Last retrieved 2013.01.28)

Remer, L. Chin, M. DeCole, P. L. Feingold, G. Halthore, R. N. Quinn, P. Rind, D. Schwartz, S. E. Streets, D. G. and Yu, H. (2008). Aerosol properties and their impacts on climate. *US climate change science program CCSP.*

Rognes, T. Mahé, F. Flouri, T. Quince C. Nichols B. (2015) vsearch: Versatile open-source tool for Metagenomics. URL https://github.com/torognes/vsearch. Last retrieved 23.11.15

Rosenstiel, T. N. Fisher, A. J. Fall, R. and Monson, R. K. (2002). Differential Accumulation of Dimethylallyl Diphosphate in Leaves and Needles of Isoprene- and Methylbutenol-Emitting and Nonemitting Species. *Plant Physiology*, **129**, 1276–1284.

Rückert, C. Birmes, F. S. Müller, C. Niewerth, H. Winkler, A. Fetzner, S. and Kalinowski, J. (2015). Complete genome sequence of *Rhodococcus erythropolis* BG43 (DSM 46869), a degrader of *Pseudomonas aeruginosa* quorum sensing signal molecules. *Journal of biotechnology*, **211**, 99-100.

Rupp DC, Anger C, Kapadia S and Totaro M (1995) Method Of Inhibiting Protozoan Growth In Eye Care Products Using A Polyvalent Cation Chelating Agent. Patent: US 5382599-A

Saitou N. and Nei M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution,* **4**, 406-425.

Sanadze, G. A. 2004. Biogenic Isoprene A Review. *Russian Journal of Plant Physiology*, **51**, 729–741.

Sanderson, M. G. Jones, C. D. Collins, W. J. Johnson, C. E. and Derwent, R. G. (2003a). Effect of Climate Change on Isoprene Emissions and Surface Ozone Levels. *Geophysical Research Letters*, **30**, 1936.

Sanderson, M. G. Jones, C. D. Collins, W. J. Johnson, C. E. and Derwent, R. G. (2003b). Effect of Climate Change on Isoprene Emissions and Surface Ozone Levels. *Geophysical Research Letters*, **30**, 10–13.

Santiago-Sotelo, P. and Ramirez-Prado, J. H. (2012). prfectBLAST: a platform-independent portable front end for the command terminal BLAST+ stand-alone suite. *BioTechniques*, **53**, 299–300.

Sasaki, K. Ohara, K. and Yazaki, K. (2005). Gene expression and characterization of isoprene synthase from *Populus alba*. *FEBS Letters*, **579**, 2514–8.

Schirmer, M. Ijaz, U. Z. D'Amore, R. Hall, N. Sloan, W. T. and Quince, C. (2015). Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Research*, **43**, 37.

Schnitzler, J. P. Graus, M. Kreuzwieser, J. Heizmann, U. Rennenberg, H. Wisthaler, A. and Hansel, A. (2004). Contribution of different carbon sources to isoprene biosynthesis in poplar leaves. *Plant Physiology*, **135**, 152–60.

Schrader, S. M. Wise, R. R. Wacholtz, W. F. Ort, D. R. and sharkey, T. D. (2004). Thylakoid membrane responses to moderately high leaf temperature in Pima cotton. *Plant, Cell and Environment*, **27**, 725–735.

Seinfeld, J. H. and Pandis, S. N. (2012). Atmospheric Chemistry and Physics: From Air Pollution to Climate Change. John Wiley & Sons.

Sequeira, L. Gaard, G. and De Zoeten, G. A. (1977). Interaction of bacteria and host cell walls: its relation to mechanisms of induced resistance. *Physiological Plant Pathology*, **10**, 43–50.

Sharkey, T. D. and Yeh, S. (2001). Isoprene emission from plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, **52**, 407–436.

Sharkey, T. D. Chen, X. and Yeh, S. (2001). Isoprene increases thermotolerance of fosmidomycin-fed leaves. *Plant Physiology*, **125**, 2001–6.

Sharkey, T. D. Loreto, F. and delwiche, C. F. (1991). High carbon dioxide and sun/shade effects on isoprene emission from oak and aspen tree leaves. *Plant, Cell and Environment*, **14**, 333–338.

Sharkey, T. D. Singsaas, E. L. Vanderveer, P. J. and Geron, C. (1996). Field measurements of isoprene emission from trees in response to temperature and light. *Tree Physiology*, **16**, 649–54.

Sharkey, T. D. Wiberley, A. E. and Donohue, A. R. (2008). isoprene emission from plants: why and how. *Annals of Botany*, **101**, 5–18.

Shaw, S. L. (2001). The Production of Non-Methane Hydrocarbons. *Marine Plankton Center for Global Change Science.*

Sheng, B. Xu, J. Zhang, Y. Jiang, T. Deng, S. Kong, J. Gao, C. Ma, C. and Xu, P. (2015). Utilization of d-Lactate as an Energy Source Supports the Growth of *Gluconobacter oxydans*. *Applied and Environmental Microbiology*, **81**, 4098–4110.

Silver, G. M. and Fall, R. (1995). Characterization of Aspen Isoprene Synthase, an Enzyme Responsible for Leaf Isoprene Emission to the Atmosphere. *Journal of Biological Chemistry,* **270**, 13010–13016.

Singsaas, E. L. and Sharkey, T. D. (2000). The effects of high temperature on isoprene synthesis in oak leaves. *Plant, Cell and Environment*, **23**, 751–757.

Singsaas, E. L. Laporte, M. M. Shi, J.-Z. Monson, R. K. Bowling, D. R. Johnson, K. Lerdau, M. Jasentuliytana, A. and Sharkey, T. D. (1999). Kinetics of leaf temperature fluctuation affect isoprene emission from red oak (*Quercus rubra*) leaves. *Tree Physiology*, **19**, 917–924.

Siwko, M. E. Marrink, S. J. de Vries, A. H. Kozubek, A. Schoot Uiterkamp, A. J. M. and Mark, A. E. (2007). Does isoprene protect plant membranes from thermal shock? A molecular dynamics study. *Biochimica et Biophysica Acta*, **1768**, 198–206.

Species, I. Singsaas, L. Lerdau, M. Winter, K. and Sharkey, T. D. (1997). Isoprene Increases Thermotolerance of Isoprene-Emitting Species*, Plant Physiology*. **115**, 1413–1420.

Srivastva, N. Shukla, A. K. Singh, R. S. Upadhyay, S. N. and Dubey, S. K. (2015). Characterization of bacterial isolates from rubber dump site and their use in biodegradation of isoprene in batch and continuous bioreactors. *Bioresource Technology*, **188**, 84-91.

Stewart E. J. (2012) Growing Unculturable Bacteria. *Journal of Bacteriology*. **194**, 4151-4160.

Tamura K. and Nei M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512-526.

Tamura K., Stecher G., Peterson D., Filipski A., and Kumar S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*. **30**, 2725-2729.

Tan, K. H. (2005). Soil Sampling, Preparation, and Analysis, Second Edition. *CRC Press*.

Telford, P. J. Lathière, J. Abraham, N. L. Archibald, A. T. Braesicke, P. Johnson, C. E. Morgenstern, O. O'Connor, F. M. Pike, R. C. Wild, O. Young, P. J. Beerling, D. J. Hewitt, C. N. and Pyle, J. (2010). Effects of climate-induced changes in isoprene emissions after the eruption of Mount Pinatubo. *Atmospheric Chemistry and Physics*, **10**, 7117–7125.

The UniProt Consortium. (2014). UniProt: a hub for protein information. *Nucleic Acids Research*, **43**, 204–212.

Thompson, I. P. Bailey, M. J. Fenlon, J. S. Fermor,T. R. Lilley, A. K. Lynch, J. M. McCormack, P. J. McQuilken, M. P. Purdy, K. J. Rainey, P. B. and Whipps J. M. (1993). Quantitative and qualitative seasonal changes in the microbial community from the phyllosphere of sugar beet (*Beta vulgaris*), *Plant and Soil,* **150**, 177-191

Tingey, D. T. Evans, R. and Gumpertz, M. (1981). Effects of environmental conditions on isoprene emission from live oak. *Planta*, **152**, 565–70.

Torsvik, V. Sørheim, R. and Goksøyr, J. (1996). Total bacterial diversity in soil and sediment communities—A review. *Journal of Industrial Microbiology*, **17**, 170–178.

Trapp, D. Cooke, K. M. Fischer, H. Bonsang, B. Zitzelsberger, R. . Seuwen, R. Schiller, C. Zenker, T. Parchatka, U. Nunes, T. V. Pio, C. A. Lewis, A. Seakins, P.and Pilling, M. (2001). Isoprene and its degradation products methyl vinyl ketone, methacrolein and formaldehyde in a eucalyptus forest during the FIELDVOC'94 campaign in Portugal. Chemosphere - *Global Change Science,* **3**, 295–307.

Trowbridge, A. M. Asensio, D. Eller, A. S. D. Way, D. A. Wilkinson, M. J. Schnitzler, J.-P. Jackson, R. B. and Monson, R. K. (2012). Contribution of Various Carbon Sources Toward Isoprene Biosynthesis in Poplar Leaves Mediated by Altered Atmospheric CO_2 Concentrations. *PLoS one*, **7**, e32387.

Turpin, B. J. (2005). Isoprene Forms Secondary Organic Aerosol through Cloud Processing. *Environmental Science and Technology,* **39**, 4441–4446.

Tyler, S.C. G. W. Braulsford, K. Yagi, K. Minami, and Cicerone R. J. (1994). Seasonal variations in methane flux and $\delta^{13}CH_4$ values for rice paddies in Japan and their implications, *Global Biogeochemical Cycles*, **8**, 1-12.

Valero-Rello, A. Hapeshi, A. Anastasi, E. Alvarez, S. Scortti, M. Meijer, W. G. & Vazquez-Boland, J. A. (2015). An invertron-like linear plasmid mediates intracellular survival and virulence in bovine isolates of *Rhodococcus equi*. Infection and immunity, **83**, 2725-2737.

Van Ginkel, C. G. de Jong, E. Tilanus, J. W. R. and de Bont, J. A. M. (1987). Microbial oxidation of isoprene, a biogenic foliage volatile and of 1,3-butadiene, an anthropogenic gas. *FEMS Microbiology Letters*, 45, 275–279.

Van Hylckama Vlieg, J. E. Kingma, J. Kruizinga, W. and Janssen, D. B. (1999). Purification of a glutathione S-transferase and a glutathione conjugate-specific dehydrogenase involved in isoprene metabolism in *Rhodococcus* sp. strain AD45. *Journal of Bacteriology*, **181**, 2094–101.

Van Hylckama Vlieg, J. E. Leemhuis, H. Spelberg, J. H. and Janssen, D. B. (2000). Characterization of the gene cluster involved in isoprene metabolism in *Rhodococcus* sp. strain AD45. *Journal of Bacteriology*, **182**, 1956–63.

Van Hylckama Vlieg, J. E. Leemhuis, H. Jeffrey, H. Spelberg, L. and Janssen, D. B. (2000). Characterization of the Gene Cluster Involved in isoprene Metabolism in *Rhodococcus* sp. Strain AD45. *Journal of Bacteriology*, **187**, 1956–1963.

Velikova, V. Varkonyi, Z. Szabo, M. Maslenkova, L. Nogues, I. Kovacs, L. Peeva, V. Busheva, M. Garab, G. Sharkey, T. D. and Loreto, F. (2011). Increased Thermostability of Thylakoid Membranes in Isoprene-Emitting Leaves Probed with Three Biophysical Techniques. *Plant Physiology*, **157**, 905–916.

Vickers, C. E. Possell, M. Cojocariu, C. I. Velikova, V. B. Laothawornkitkul, J. Ryan, A. Mullineaux, P. M. and Hewitt, C. N. (2009). Isoprene synthesis

protects transgenic tobacco plants from oxidative stress. *Plant, Cell & Environment*, **32**, 520–31.

Vickers, C. E. Possell, M. Nicholas Hewitt, C. and Mullineaux, P. M. (2010). Genetic structure and regulation of isoprene synthase in Poplar *Populus* spp.. *Plant Molecular Biology,* **73**, 547–58.

Vorholt J. A. (2012). Microbial life in the phyllosphere. *Nature reviews microbiology*. **10**. 828-840

Wagner (2015). vegan: Community Ecology Package. *R package* version 2.2-1.

Wagner, P. and Kuttler, W. (2014). Biogenic and anthropogenic isoprene in the near-surface urban atmosphere—A case study in Essen, Germany. *Science of the Total Environment*, **475**, 104-115.

Wang, Q. Garrity, G. M. Tiedje, J. M. and Cole, J. R. (2007). Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology*, **73**, 5261-7.

Wang, Y. I. Naumann, U. Wright, S. T. and Warton, D. I. (2012). mvabund–an R package for model-based analysis of multivariate abundance data. *Methods in Ecology and Evolution*, ***3***, 471-474.

Whitman, W. B. Coleman, D. C. and Wiebe, W. J. (1998). Perspective Prokaryotes : The unseen majority, *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 6578–6583.

Wiberley, A. E. Donohue, A. R. Westphal, M. M. and Sharkey, T. D. (2009). Regulation of isoprene emission from poplar leaves throughout a day. *Plant, Cell & Environment*, **32**, 939–47.

Wildermuth, M. C. and Fall, R. (1996). Light-Dependent Isoprene Emission Characterization of a Thylakoid-Bound Isoprene Synthase in *Salix discolor* Chloroplasts. *Plant Physiology*, **112**, 171–182.

Wilson, M. and Lindow, S. E. (1994). Coexistence among Epiphytic Bacterial Populations Mediated through Nutritional Resource Partitioning Coexistence, *American Society for Microbiology,* **60**, 4468–4477.

Wolfertz, M. Sharkey, T. D. Boland, W. and Kühnemann, F. (2004). Rapid regulation of the methylerythritol 4-phosphate pathway during isoprene synthesis. *Plant Physiology*, **135**, 1939–1945.

World Health Organisation, (2014), Burden of disease from Household Air Pollution for 2012

World Health Organisation. (2000). Air Quality Guidelines for Europe. *WHO Regional Publications*, European Series.

Xie S, Sun W and Luo C (2011) Novel aerobic benzene degrading microorganisms identified in three soils by stable isotope probing. *Biodegradation*. **22**,71-81.

Xun, L. Topp, E. and Orser, C. S. (1992). Purification and characterization of a tetrachloro-p-hydroquinone reductive dehalogenase from a *Flavobacterium* sp. *Journal of Bacteriology*, **174**, 8003–8007.

2.SI: Chapter 2 Supplementary Information

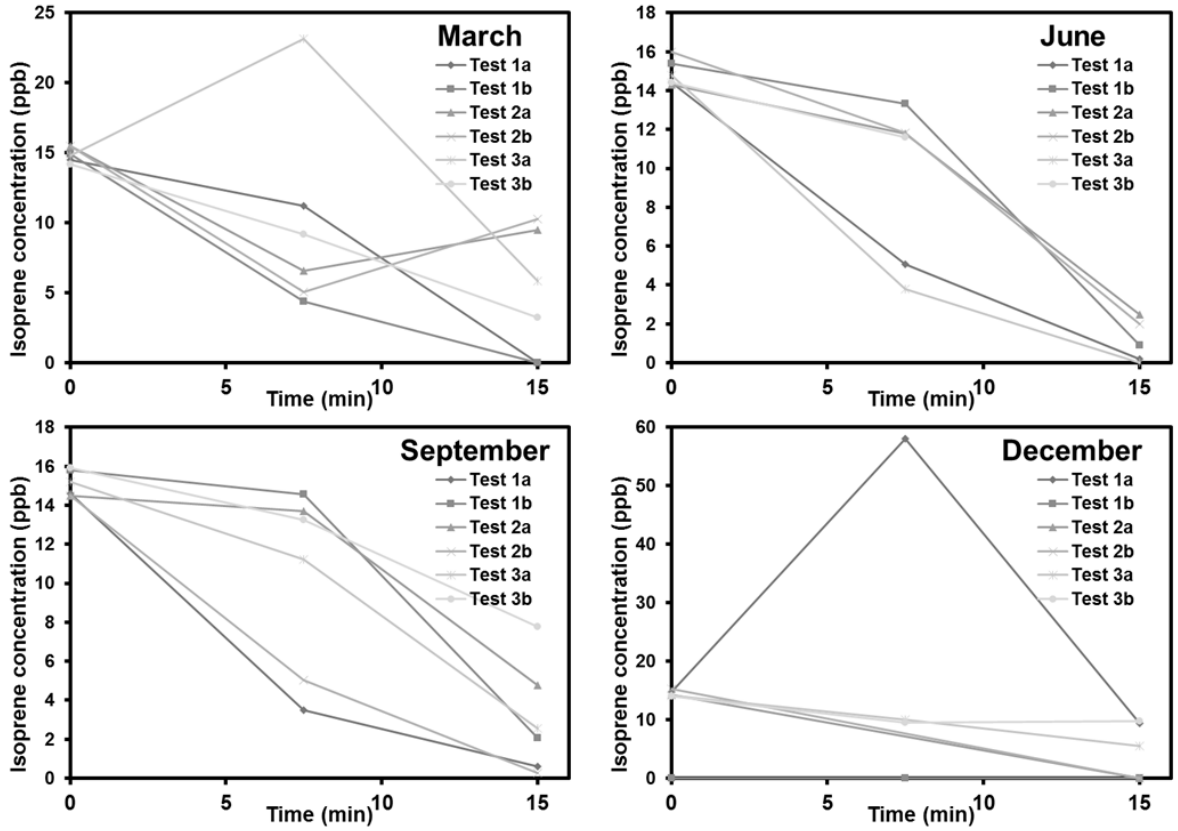
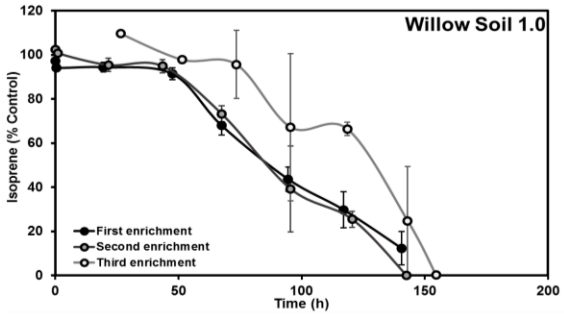
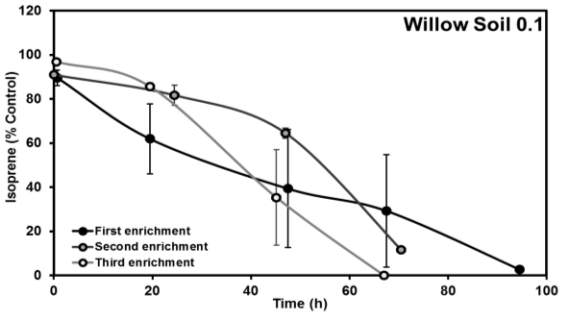
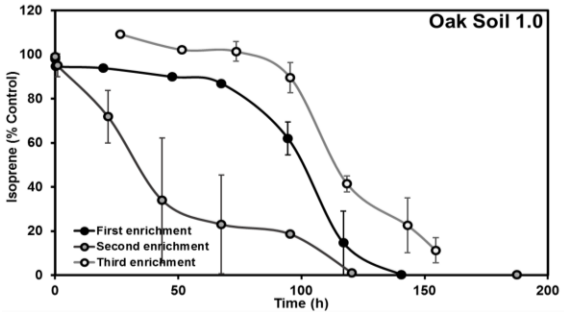
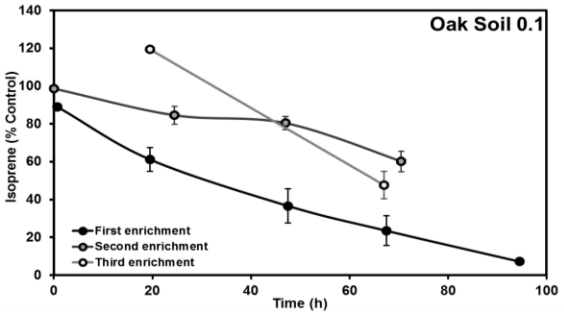


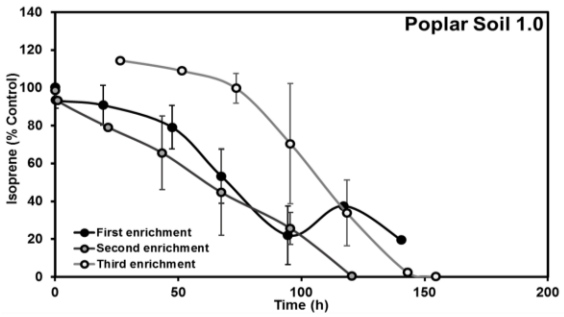
Figure 2.SI.1 Isoprene concentration over time for 2 l chambers implanted into soil and spiked with ~15 ppb isoprene as well Decafluoropentane for different time-points in 2015 (24/3/15, 18/6/15, 29/9/15, 10/12/15). Concentration calculated as the change in Isoprene:DFP ratio, net of in-group control ratio change, against a standard curve. Chambers in three groups, with two test chambers per group. n = 6 (failed data included).

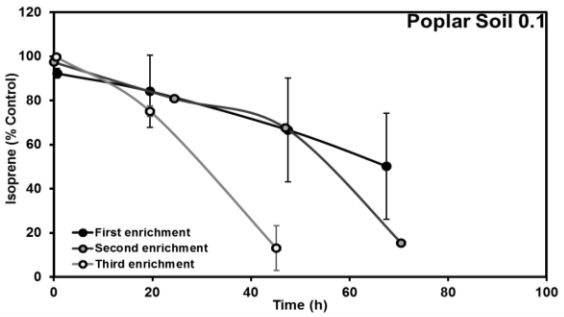

Willow Soil 1.0
Isoprene (% Control)
Time (h)
First enrichment
Second enrichment
Third enrichment

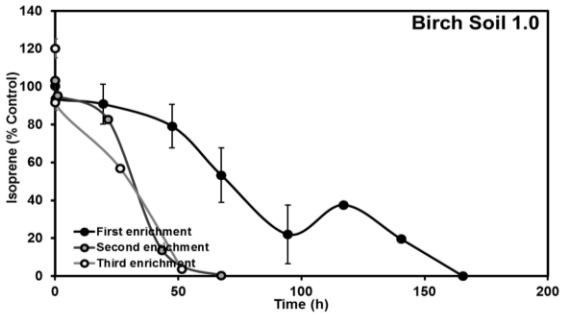

Willow Soil 0.1
Isoprene (% Control)
Time (h)
First enrichment
Second enrichment
Third enrichment

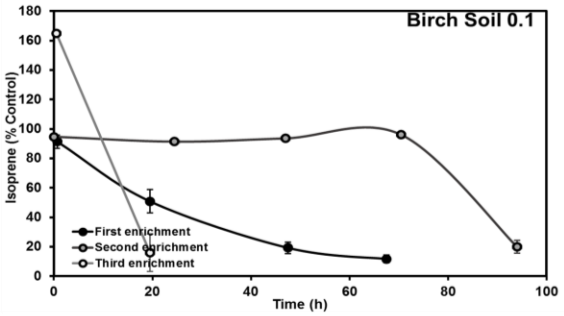

Oak Soil 1.0
Isoprene (% Control)
Time (h)
First enrichment
Second enrichment
Third enrichment

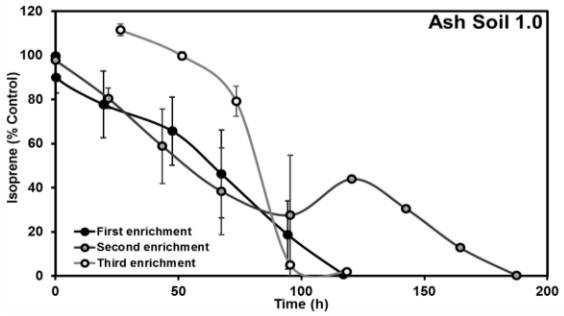

Oak Soil 0.1
Isoprene (% Control)
Time (h)
First enrichment
Second enrichment
Third enrichment

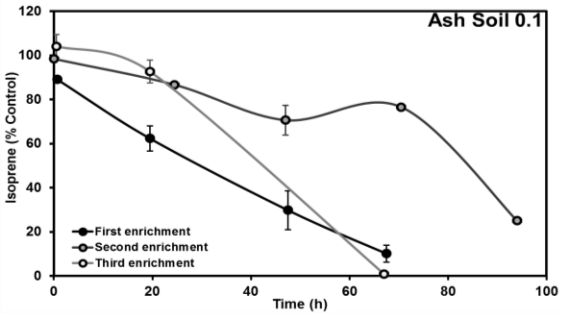

Poplar Soil 1.0
Isoprene (% Control)
Time (h)
First enrichment
Second enrichment
Third enrichment

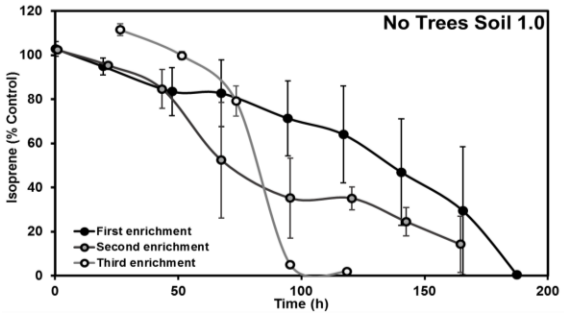

Poplar Soil 0.1
Isoprene (% Control)
Time (h)
First enrichment
Second enrichment
Third enrichment


Birch Soil 1.0
Isoprene (% Control)
Time (h)
First enrichment
Second enrichment
Third enrichment


Birch Soil 0.1
Isoprene (% Control)
Time (h)
First enrichment
Second enrichment
Third enrichment


Ash Soil 1.0
Isoprene (% Control)
Time (h)
First enrichment
Second enrichment
Third enrichment


Ash Soil 0.1
Isoprene (% Control)
Time (h)
First enrichment
Second enrichment
Third enrichment


No Trees Soil 1.0
Isoprene (% Control)
Time (h)
First enrichment
Second enrichment
Third enrichment

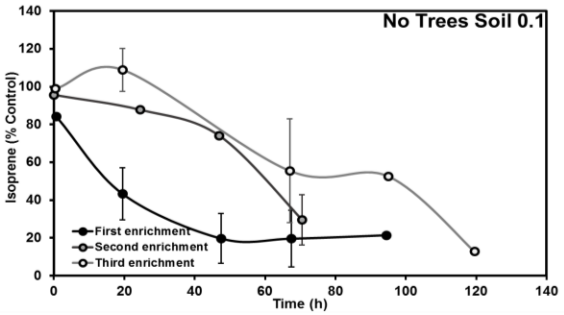

No Trees Soil 0.1
Isoprene (% Control)
Time (h)
First enrichment
Second enrichment
Third enrichment

Figure 2.SI.2 Isoprene concentration measurements against time, transformed as percentage of average abiotic controls for 1.0 cm^3 (left) and 0.1 cm^3 (right) 30°C saturated isoprene headspace addition, for six soil origins from under the canopy of Willow, Oak, Poplar, Birch, Ash, and No-Trees (top to bottom), for three sequential soil enrichments with isoprene, (enrichment 1 = black, enrichment 2 = grey, enrichment 3 = white) in minimal media. Third level of enrichment sometimes out of sync (starting late as second level finish times were different), initial n = 3, error bars ±SE.

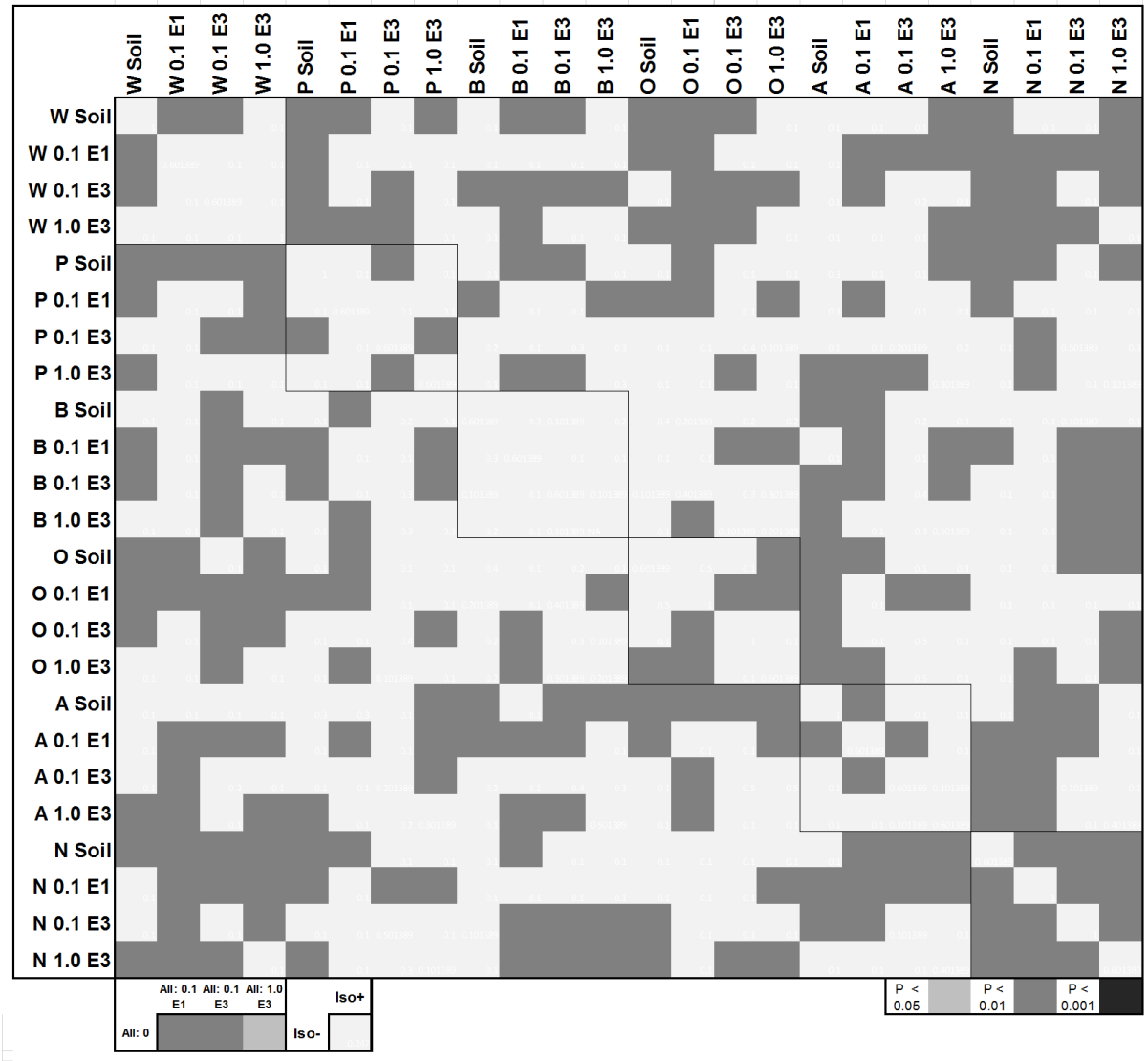


Figure 2.SI.3 Presence and level of significant differences. Significance levels measured by permutational manovas (R: Vegan: adonis) for Willow (W), Poplar (P), Birch (B), Oak (O), Ash (A) and No-Trees (N) for Soil, and Isoprene enrichments with 0.1 and 1.0 ml 30°C saturated isoprene headspace additions, at the first (E1) and third (E3) sequential enrichment, against each other condition. n=3 (B 1.0 E3 n=2), combined samples at different enrichment levels, and samples from combined isoprene producing trees against from non-producing trees. White represents no significant differences.

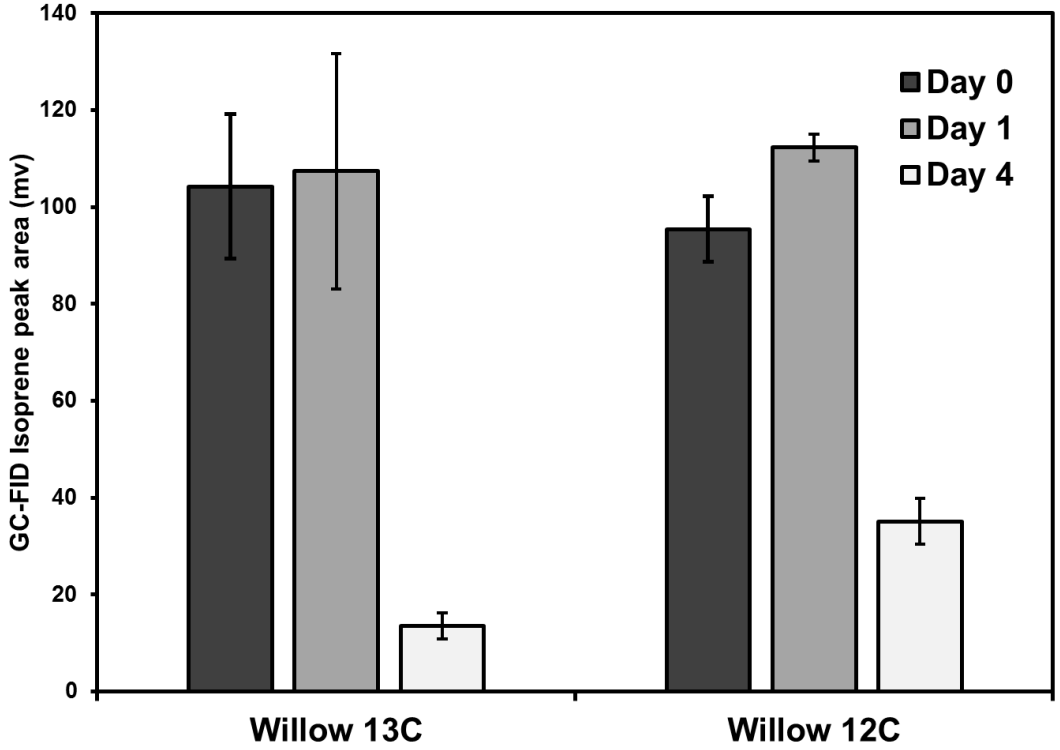


Figure 2.SI.4 Isoprene concentration for Willow soil enrichments with ^{13}C and ^{12}C isoprene for the initial isoprene level (Day 0, dark grey), Day 1 (mid-grey) and Day 4 (white). n=3, error bars ±SE.

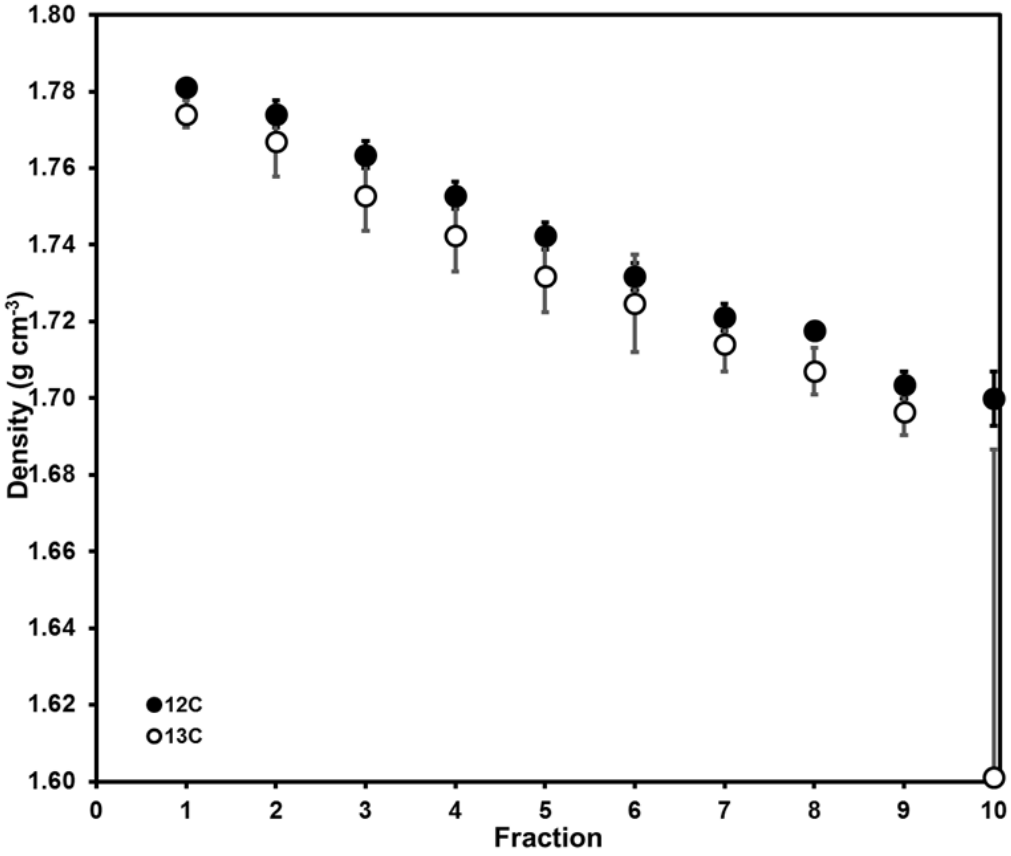


Figure 2.SI.5 Density for fractions from CsCl centrifugation after incubation of soil from near a Willow tree with ^{13}C (white) and ^{12}C DNA-SIP (black) isoprene. n=3 Error bars are ±SE.

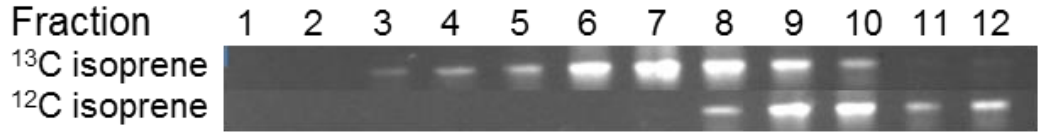


Figure 2.SI.6 Example replicate showing shift of visible presence of end-point PCR product following PCR of the 16s region for DNA-SIP fractions 1 (heaviest) to 12 (lighter) for ^{13}C and ^{12}C isoprene enrichments of soil.

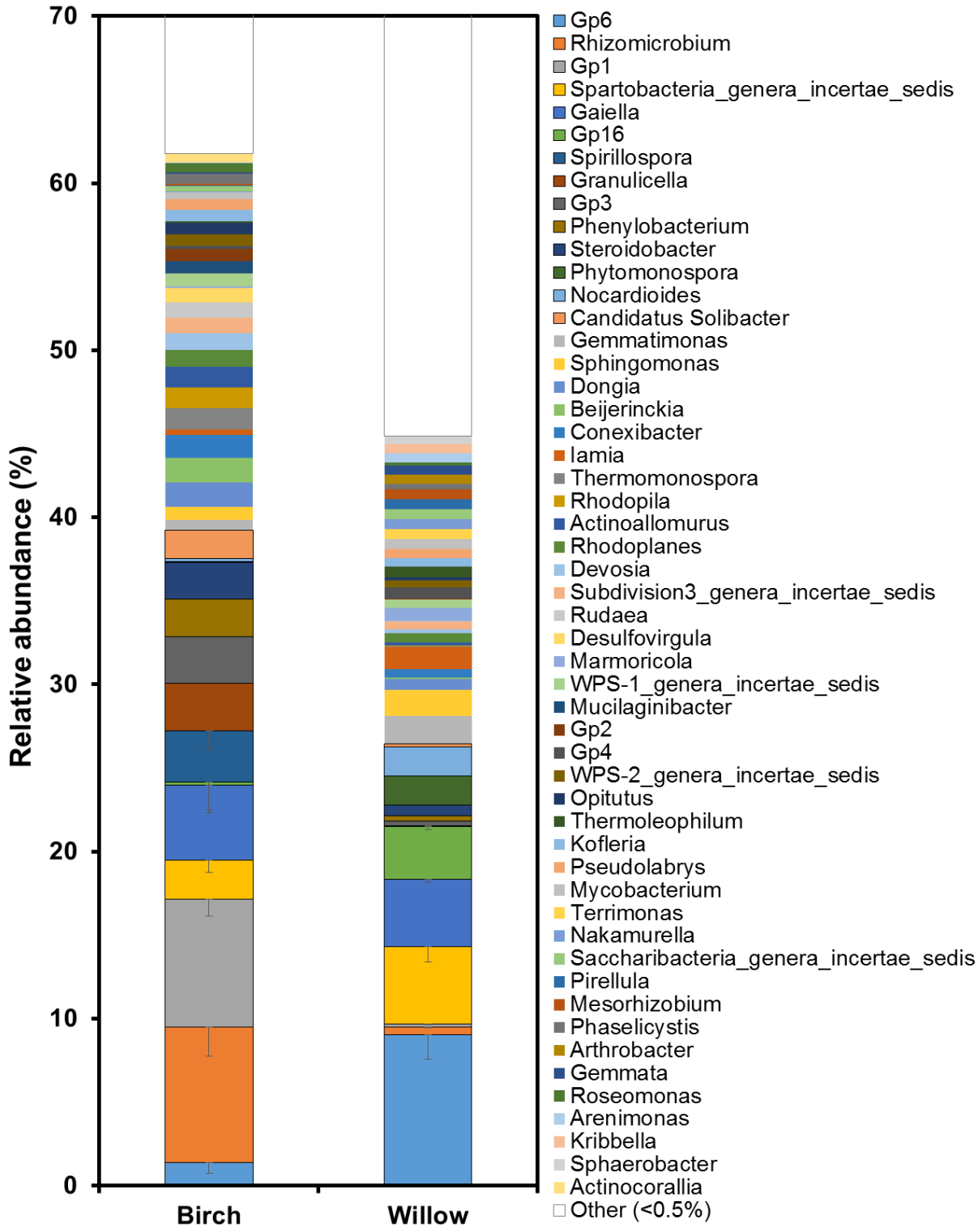


Figure 2.SI.7 Initial soil communities at RDP fixrank resolution for low concentration and stable-isotope probing experiments after MiSeq amplicon sequencing of bacterial 16s rDNA genes. n=3, Error bars are ±SE and included only for communities above 3% relative abundance.

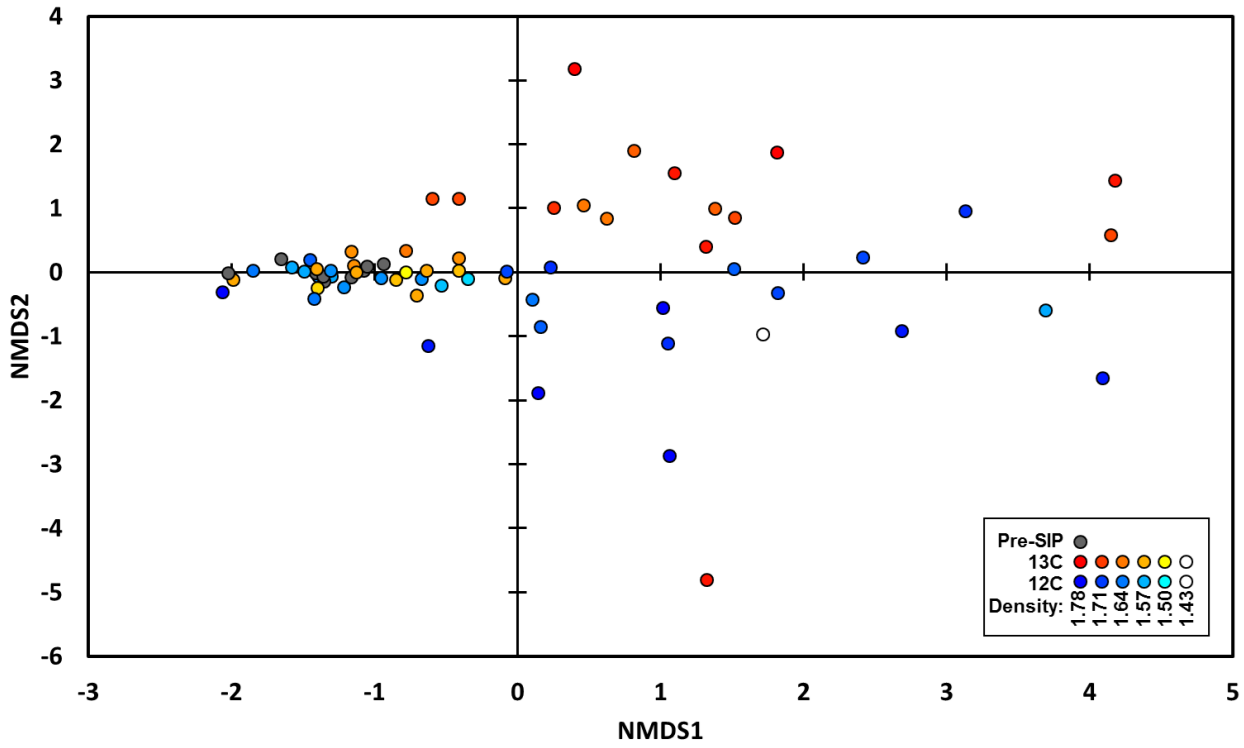


Figure 2.SI.8 NMDS of Stable isotope probing of soil with heavy (^{13}C) isoprene, normal isoprene (^{12}C) and pre-enriched soil. ^{13}C Red (high density) to Yellow (low density), n=30, ^{12}C Blue (high density) to Turquoise (low density), n=30, Pre-enriched soil in grey, n=10. Soil from beneath Willow. Species n=30877 Performed using metaMDS implementation of isoMDS in Vegan package in R; run 2207, stress 0.1401088.

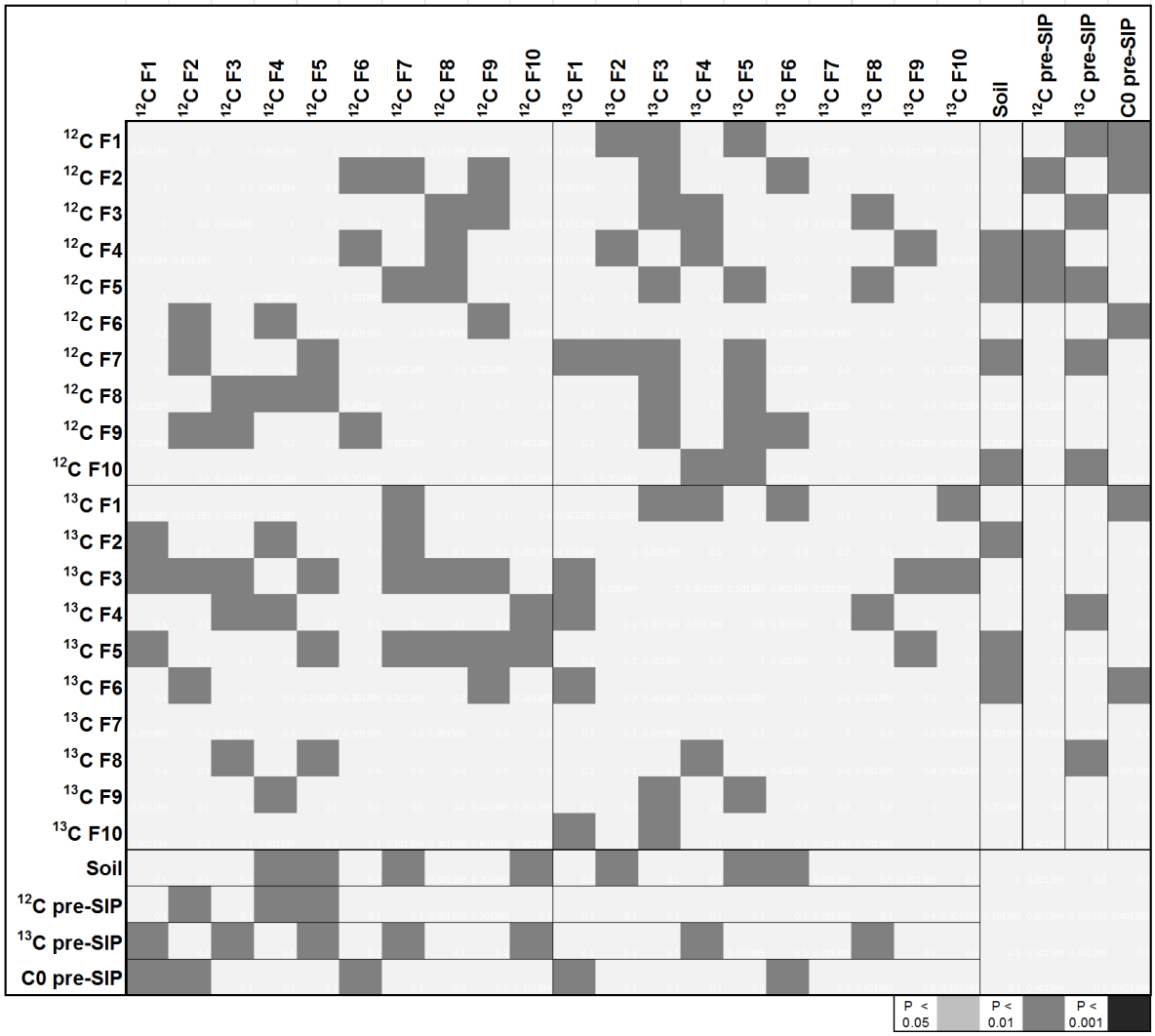


Figure 2.SI.9 Significance distribution of ^{12}C and ^{13}C ultracentrifugation fractions 1 to 10 (^{12}C F1 to ^{12}C F10; ^{13}C F1 to ^{13}C F10) from DNA-SIP of soil with isoprene and heavy isoprene, the pre-enriched soil (Soil), and the pre-centrifugation communities (pre-SIP) for ^{12}C, ^{13}C and no isoprene (C0) incubation communities measured by permutational manovas (R: Vegan: adonis), n=3. White represents no significant difference.

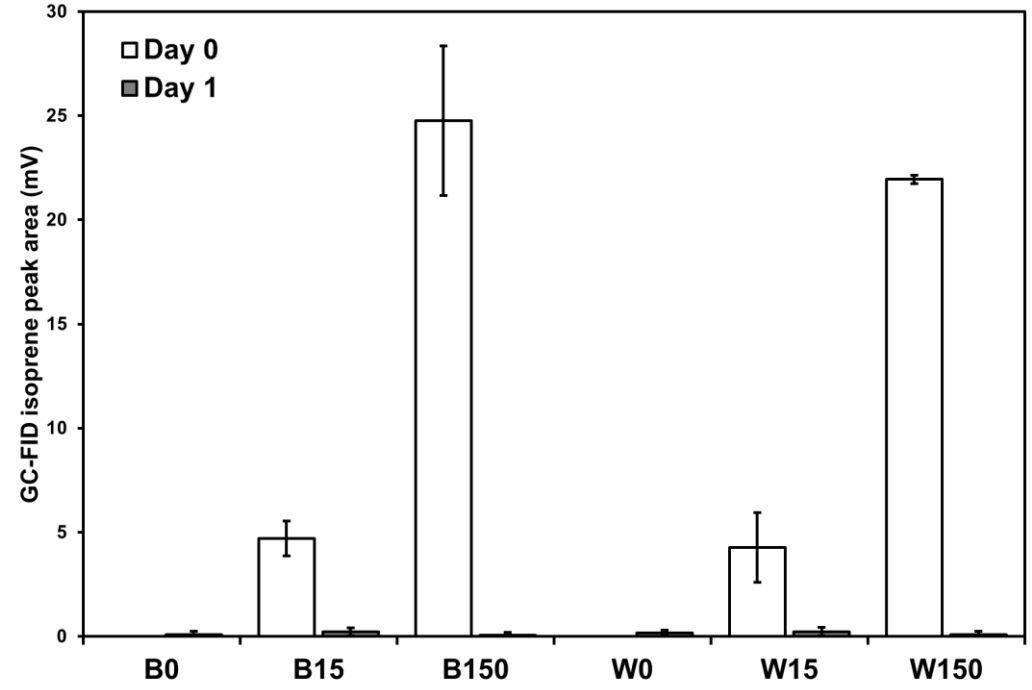


Figure 2.SI.10 example of daily isoprene decrease under low concentrations of isoprene, isoprene was added to cultures and tested (white bars), and retested after 1 day (grey bars) before refilling; with Birch soil (B), Willow soil (W), 0, 15 and 150 ppb isoprene, n=3, Error bars are ± SE, data = 2nd-3rd June.

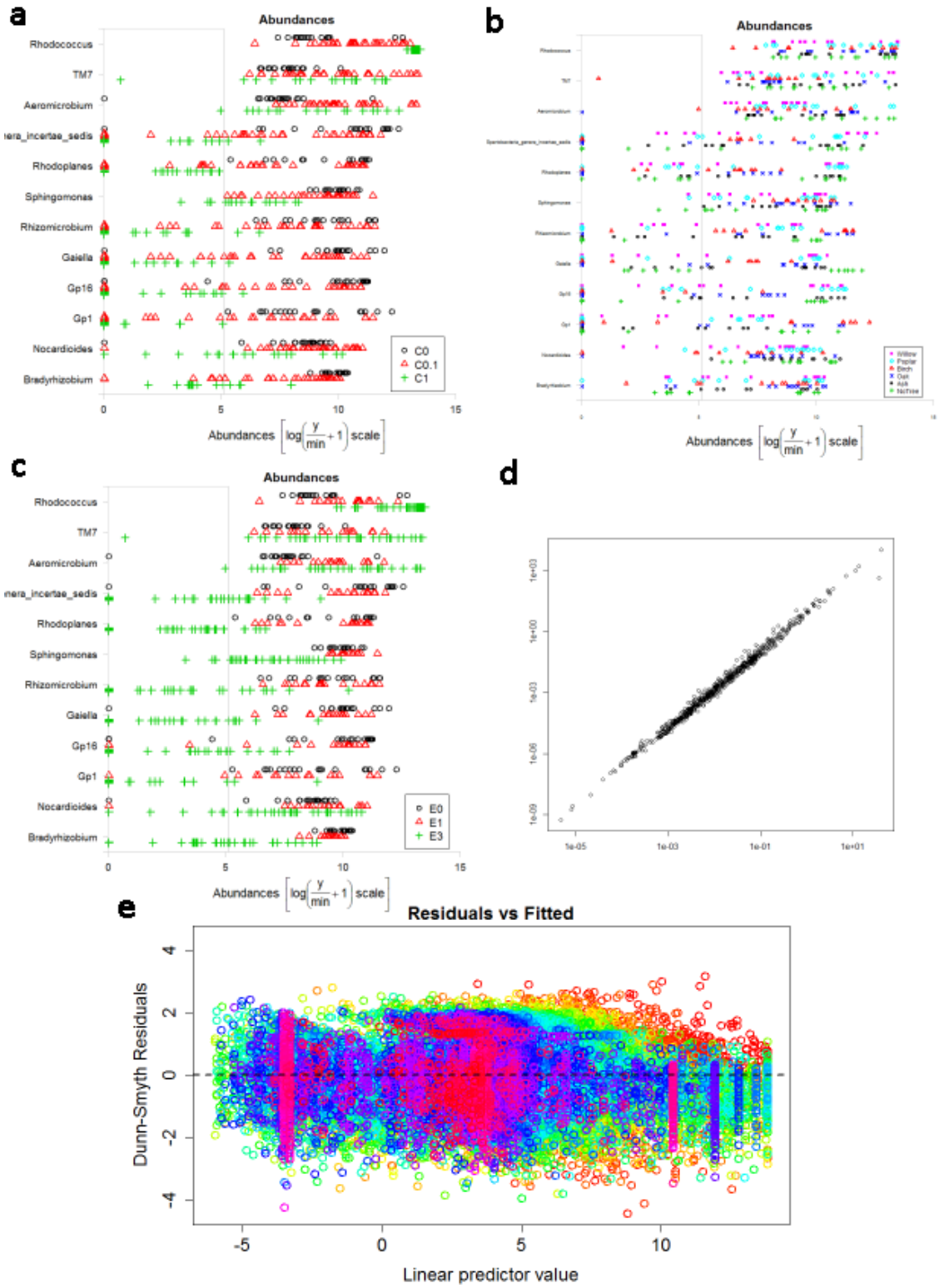


Figure 2.SI.11 Serial enrichment experiments mvabund output: Distribution of abundaces for most abundant bacteria for (a) different concentrations of isoprene (C0 = no isoprene, C0.1 = $7.2*10^5$ ppb isoprene, C01 = $7.2*10^6$ ppb isoprene) (b) different tree types overshadowing the sample location and (c) different levels of enrichment (E1 is the first enrichment, E3 is the third sequential enrichment), with (d) a plot of the mean variance of the final model and (e) a plot of the Dunn-Smyth Residuals against the linear predictor values.

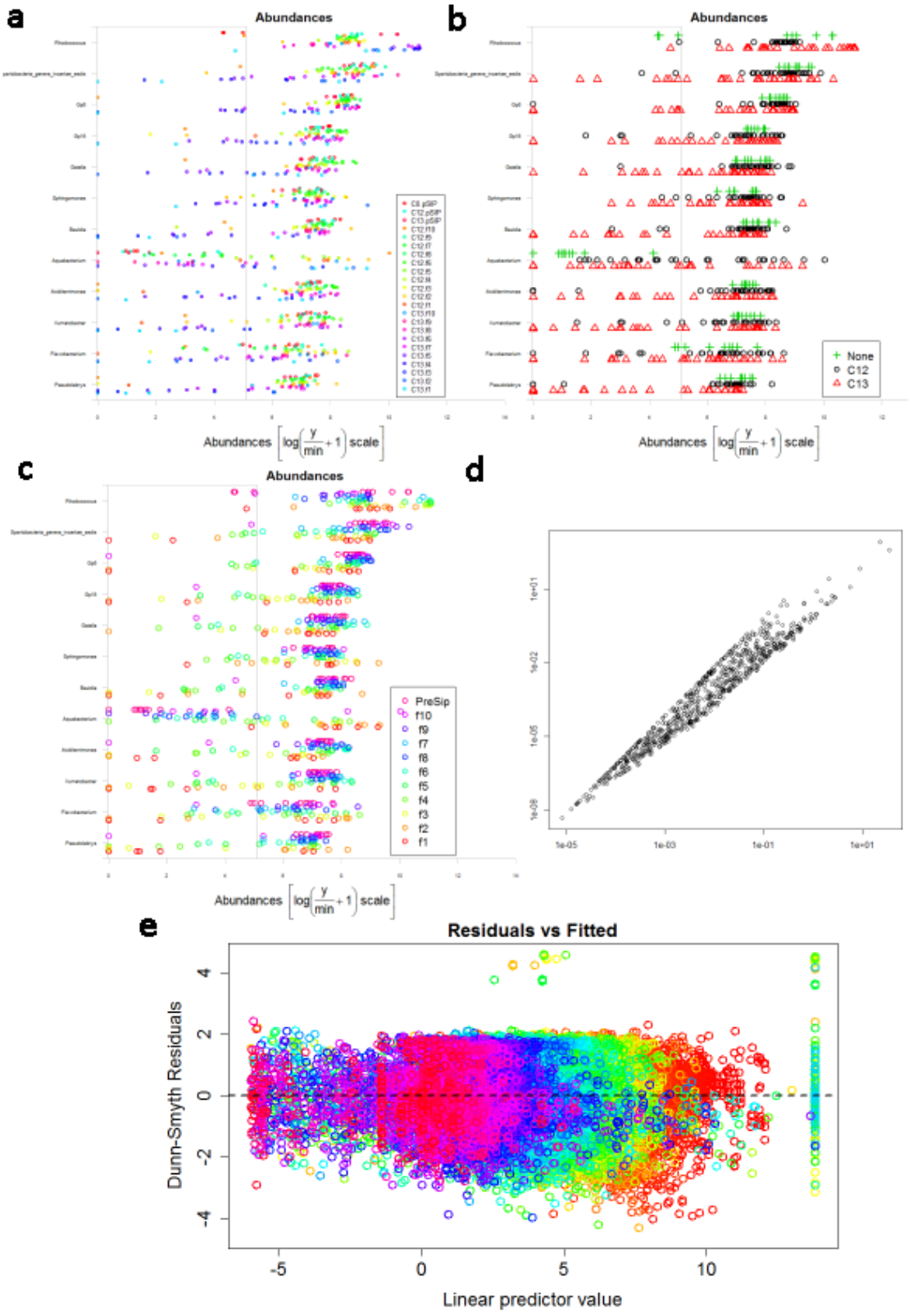


Figure 2.SI.12 DNA-SIP experiments mvabund output: a) Distribution of abundaces for most abundant bacteria for (a) each fraction f1-f10 (f1 being the heaviest) for the starting (pSIP), and final communities for the C12 and C13 experiments (b) different communities within the Carbon-12, Carbon-13 and untreated groups (c) different fractions following density

gradient centrifugation (f1 is the most dense, f10 the least) (d) a plot of the mean variance of the final model and (e) a plot of the Dunn-Smyth Residuals against the linear predictor values.

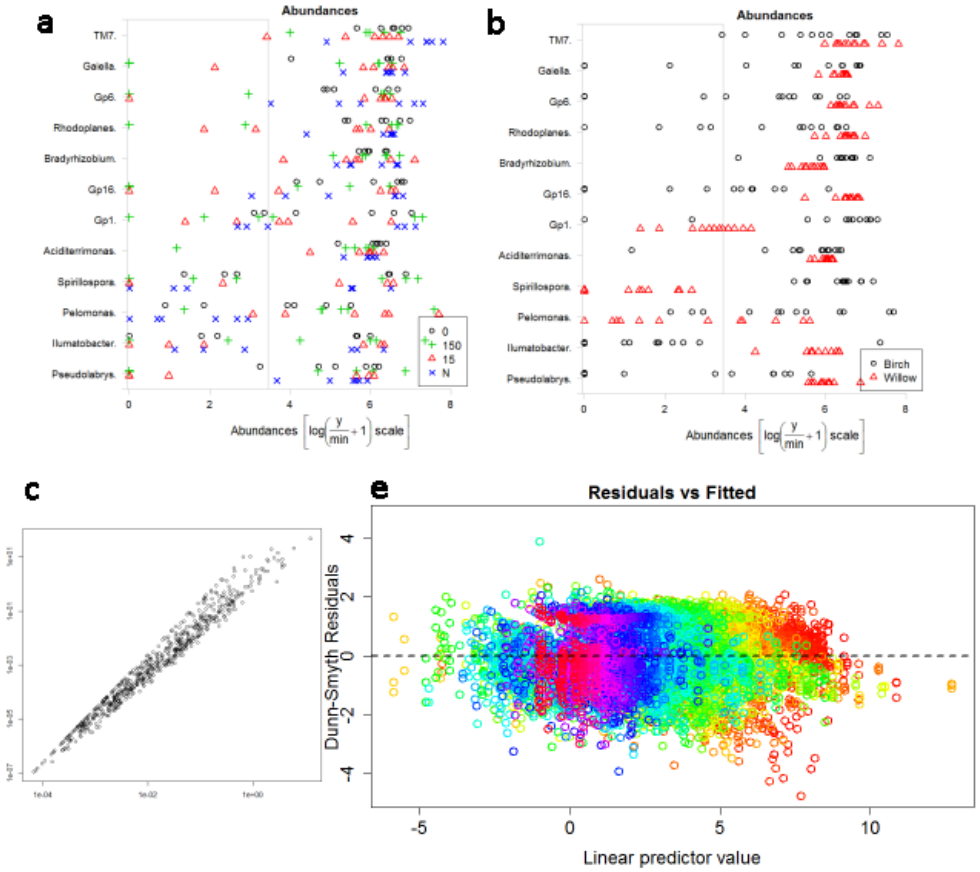


Figure 2.SI.13 Low Concentration experiments mvabund output: Distribution of abundaces for most abundant bacteria for (a) different concentrations of isoprene N = starting community, 150=150 ppb isoprene, 15 = 15 ppb isoprene, 0 = no isoprene control (b) different tree types overshadowing the sample location, with (c) a plot of the mean variance of the final model and (d) a plot of the Dunn-Smyth Residuals against the linear predictor values.

Figure 2.SI.14 *Aeromicrobium* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 1.34 P=0.24), was affected by the isoprene concentration ($F_{8,62}$ = 3.19 P=0.01),

but was not affected by the level of enrichment ($F_{8,62}$ = 1.44 P=0.2), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Aeromicrobium* defined by RDP classification at deepest assignment.

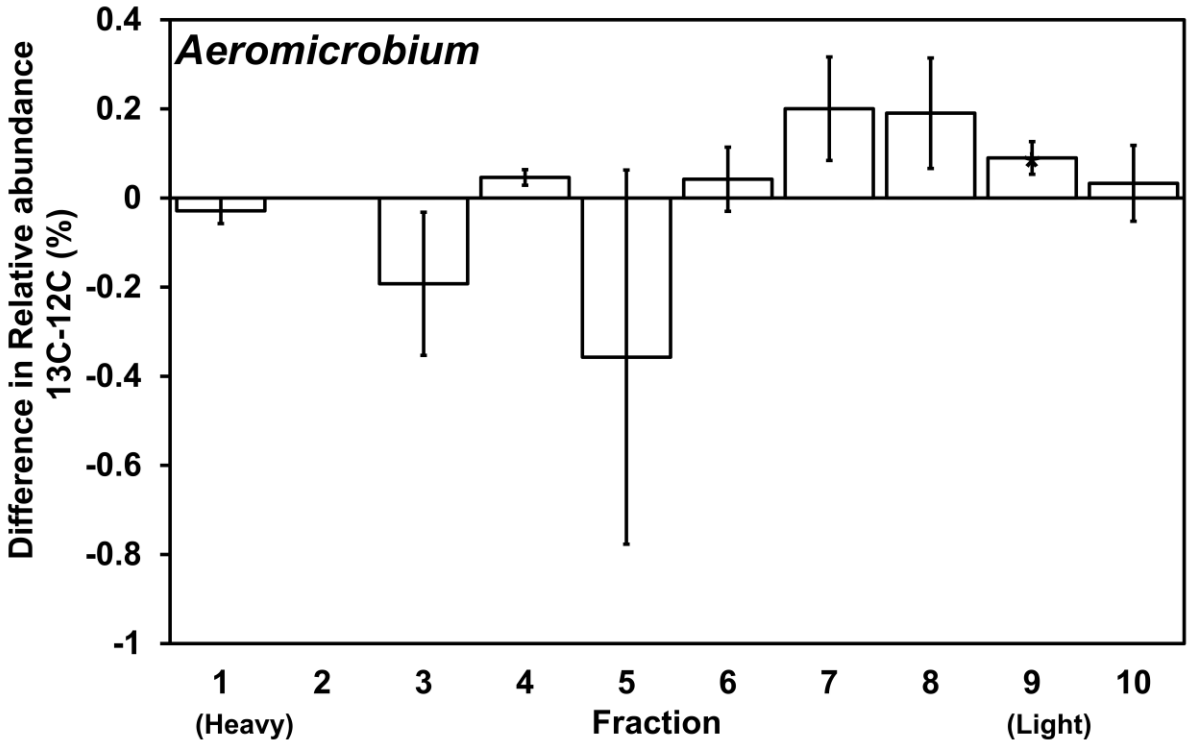


Figure 2.SI.15 *Aeromicrobium* relative abundances in ^{13}C fractions net of *Aeromicrobium* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Aeromicrobium* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by the carbon type ($F_{3,65}$ = 2.18 P=0.1), and was not affected by the density ($F_{3,65}$ = 1.12 P=0.35), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

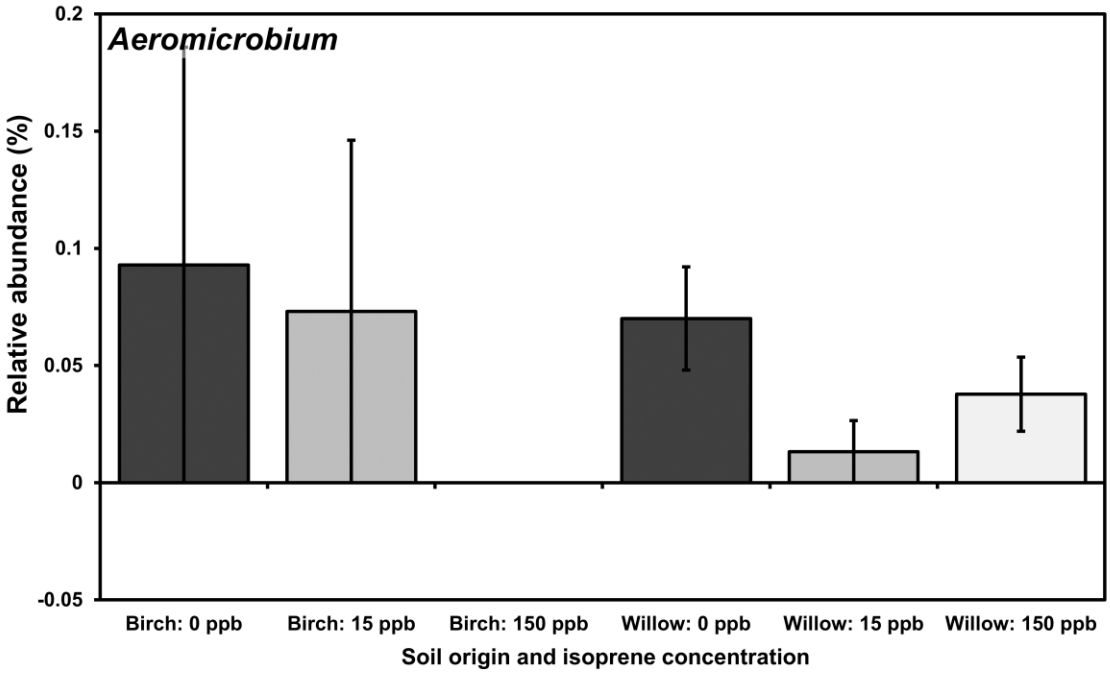


Figure 2.SI.16 *Aeromicrobium* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Aeromicrobium* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that *abundance* was not affected by the isoprene concentration ($F_{8,15}$ = 1.43 P=0.26) nor tree type ($F_{8,15}$ = 1.09 P=0.42)

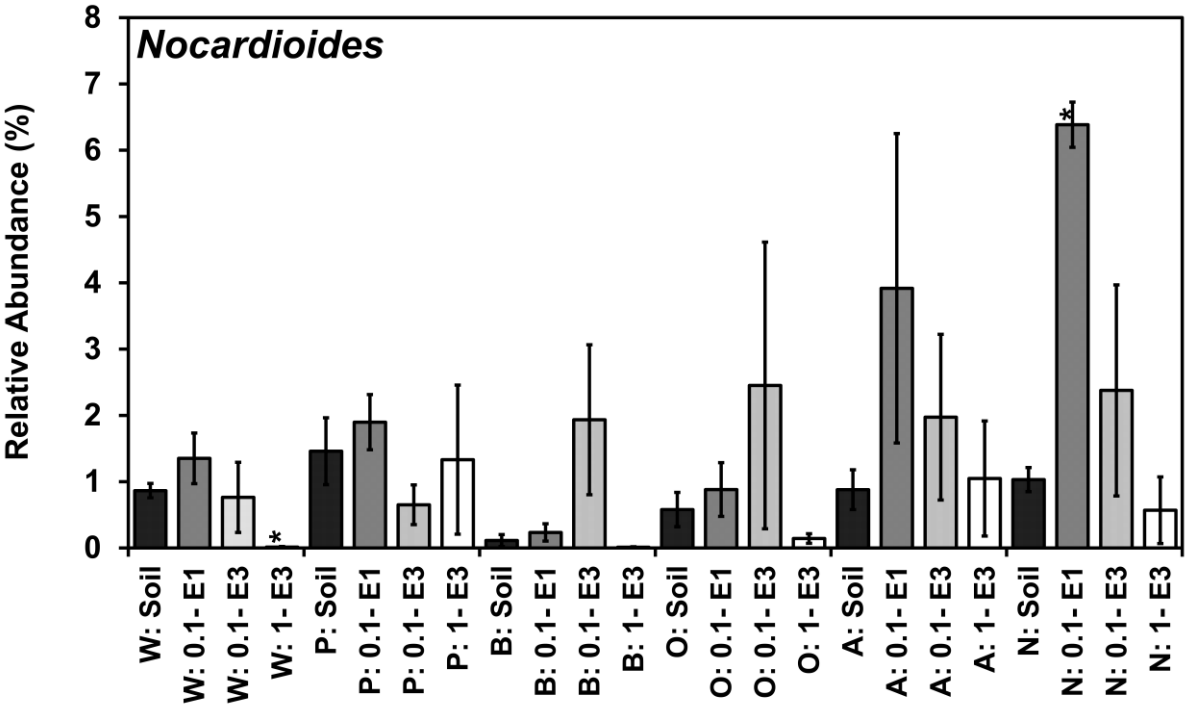


Figure 2.SI.17 relative abundances in sequential enrichment of soil with

isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3.A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was affected by concentration ($F_{8,62}$ = 3.19 P=0.04), but was not affected by sample location ($F_{8,62}$ = 1.42 P=0.21), or by enrichment level ($F_{8,62}$ = 0.84 P=0.58), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Nocardioides* defined by RDP classification at deepest assignment.

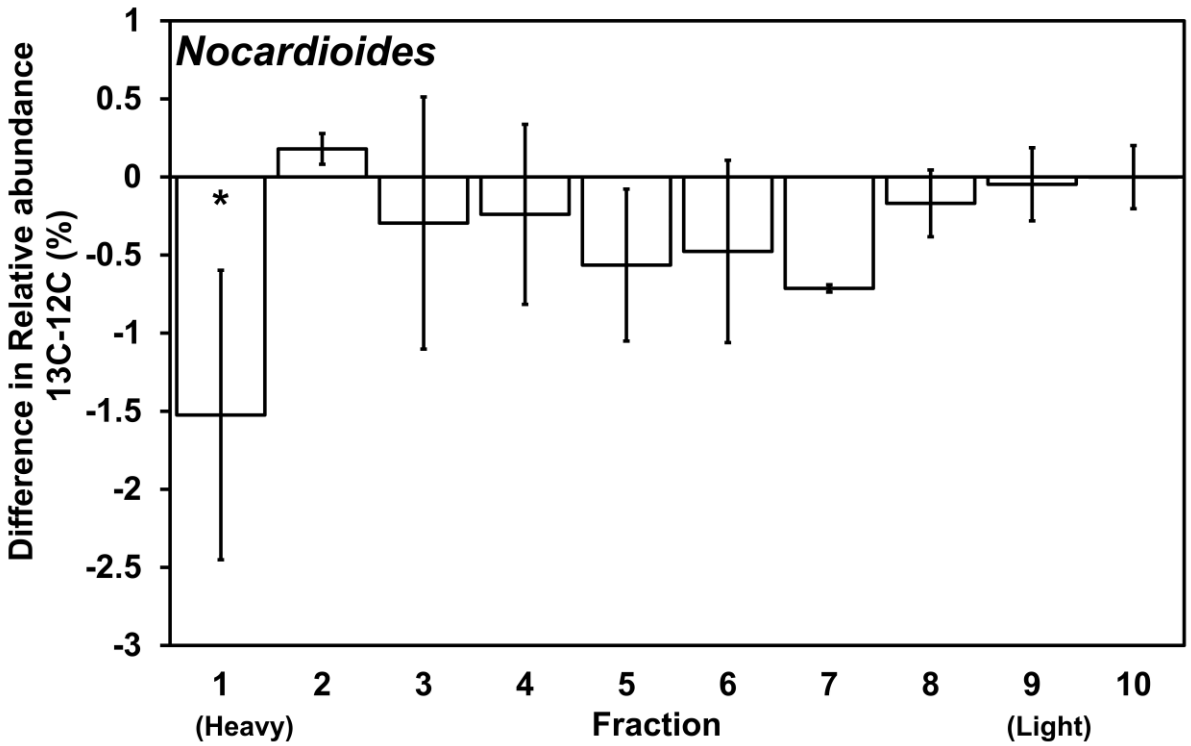


Figure 2.SI.18 *Nocardioides* relative abundances in ^{13}C fractions net of *Nocardioides* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Nocardioides* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not significantly affected by the density ($F_{3,65}$ = 1.49 P=0.23), and was not affected by the carbon type ($F_{3,65}$ = 0.89 P=0.45), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only)

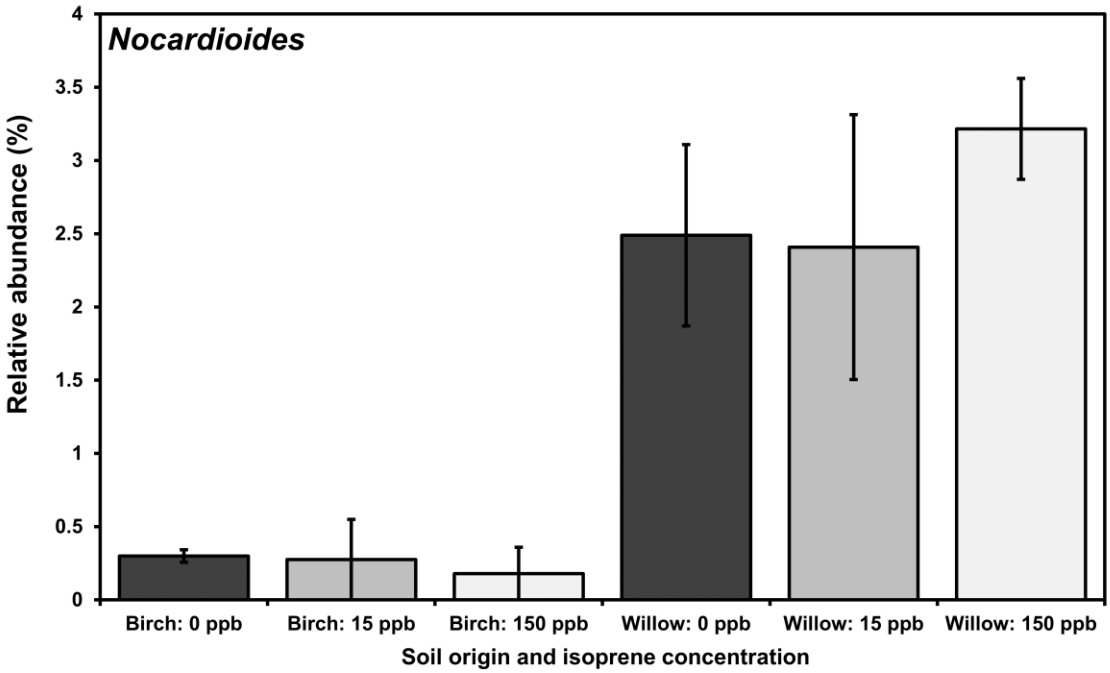


Figure 2.SI.19 *Nocardioides* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Nocardioides* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 0.3 P=0.96) and was affected by the tree type ($F_{8,15}$ = 3.37 P=0.02).

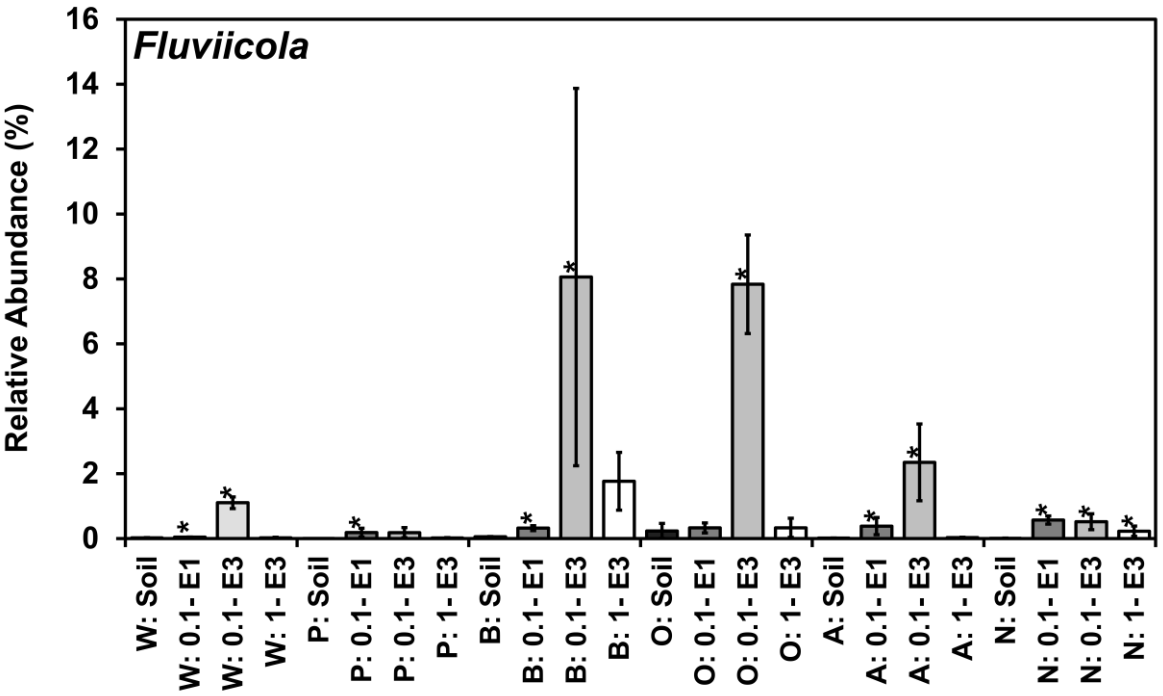


Figure 2.SI.20 Fluviicola relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment

was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was affected by sample location ($F_{8,62}$ = 2.1 P=0.05), was affected by concentration ($F_{8,62}$ = 3.81 P= < 0.001) and was was affected by enrichment level ($F_{8,62}$ = 3.19 P=0.03), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Fluviicola* defined by RDP classification at deepest assignment.

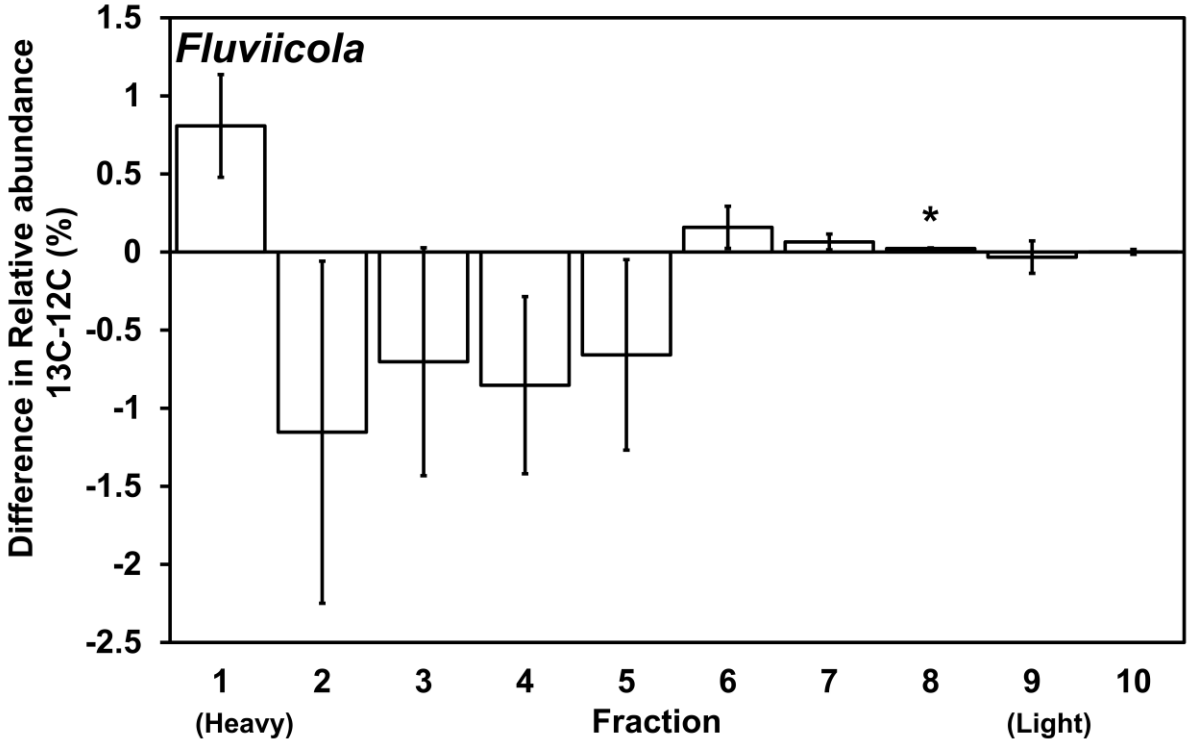


Figure 2.SI.21 *Fluviicola* relative abundances in ^{13}C fractions net of *Fluviicola* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Fluviicola* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by the density ($F_{3,65}$ = 4.4 P=0.01), and was not affected by the carbon type ($F_{3,65}$ = 0.75 P=0.53), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

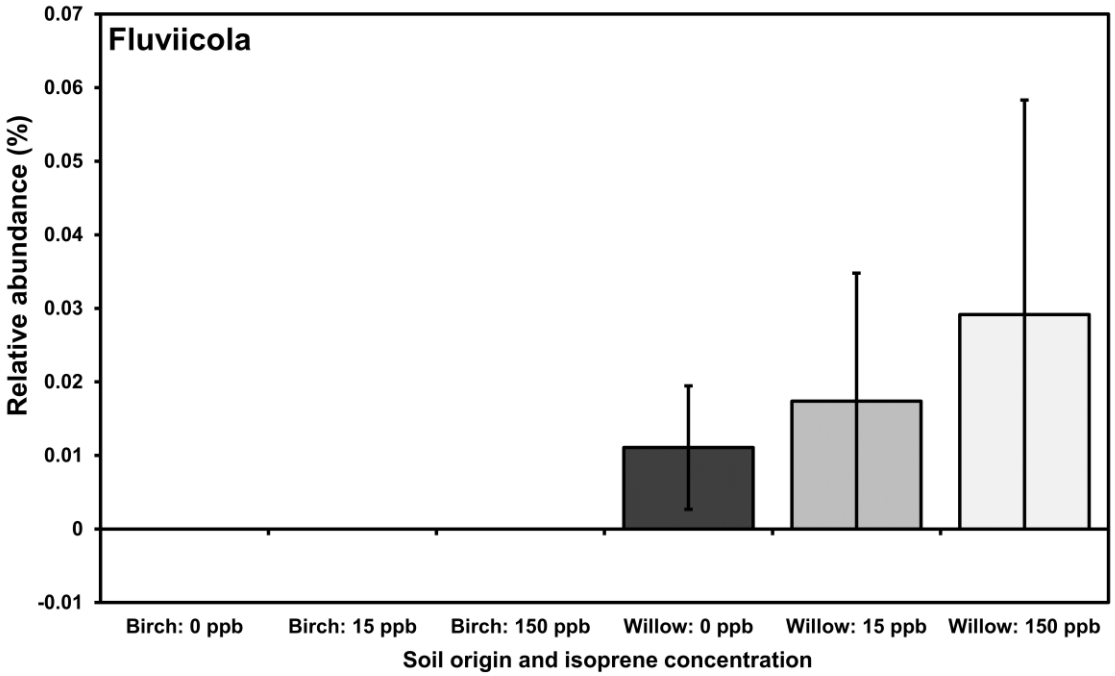


Figure 2.SI.22 *Fluviicola* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Fluviicola* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 0.01 P=1) and was not affected by the tree type ($F_{8,15}$ = 0.83 P=0.59).

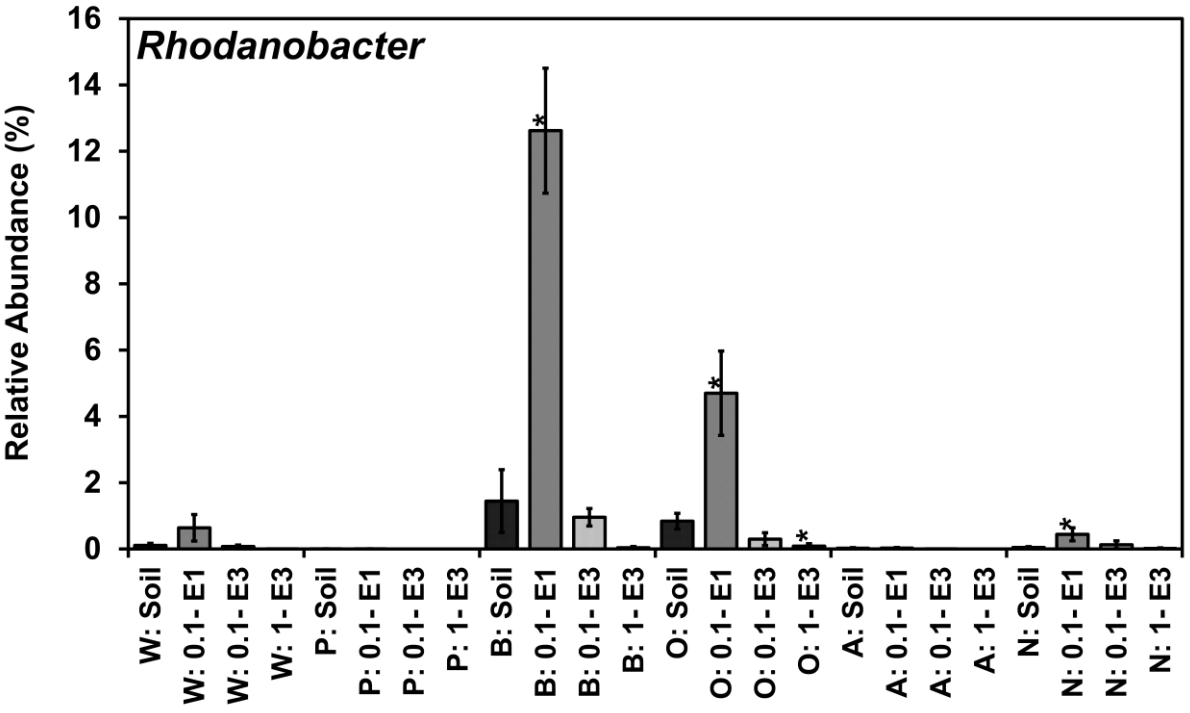


Figure 2.SI.23 *Rhodanobacter* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 0.82 P=0.6), not affected by concentration ($F_{8,62}$ = 0.32 P=0.96), and was affected by enrichment level ($F_{8,62}$ = 3.19 P=0.01), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Rhodanobacter* defined by RDP classification at deepest assignment.

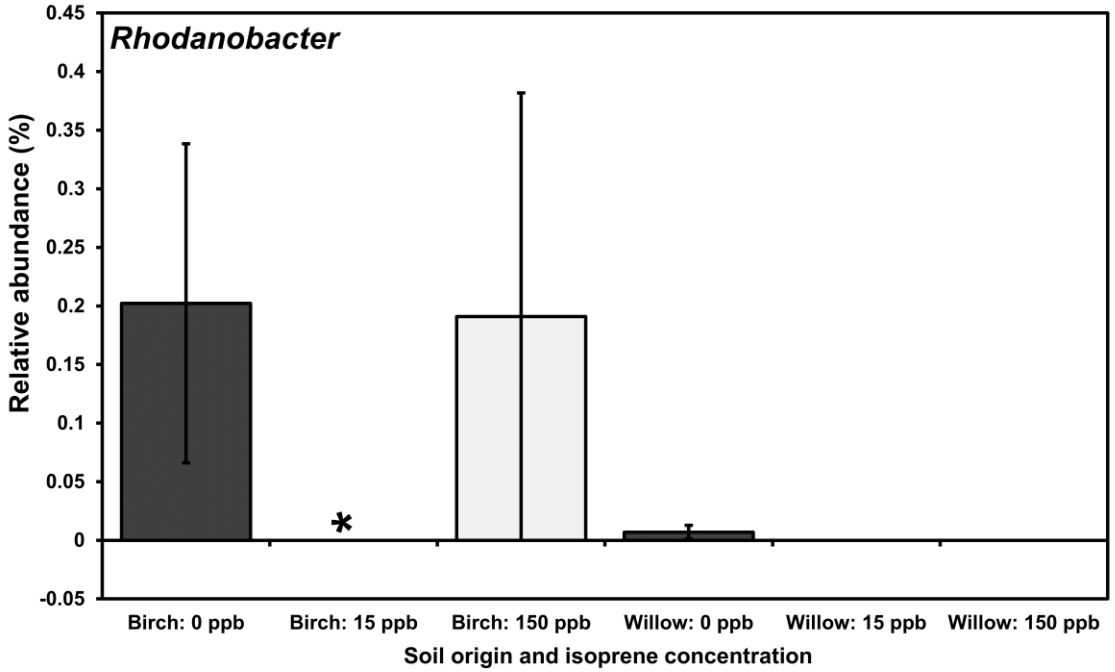


Figure 2.SI.24 *Rhodanobacter* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 1.3 P=0.32) and was not affected by the tree type ($F_{8,15}$ = 2.35 P=0.07).

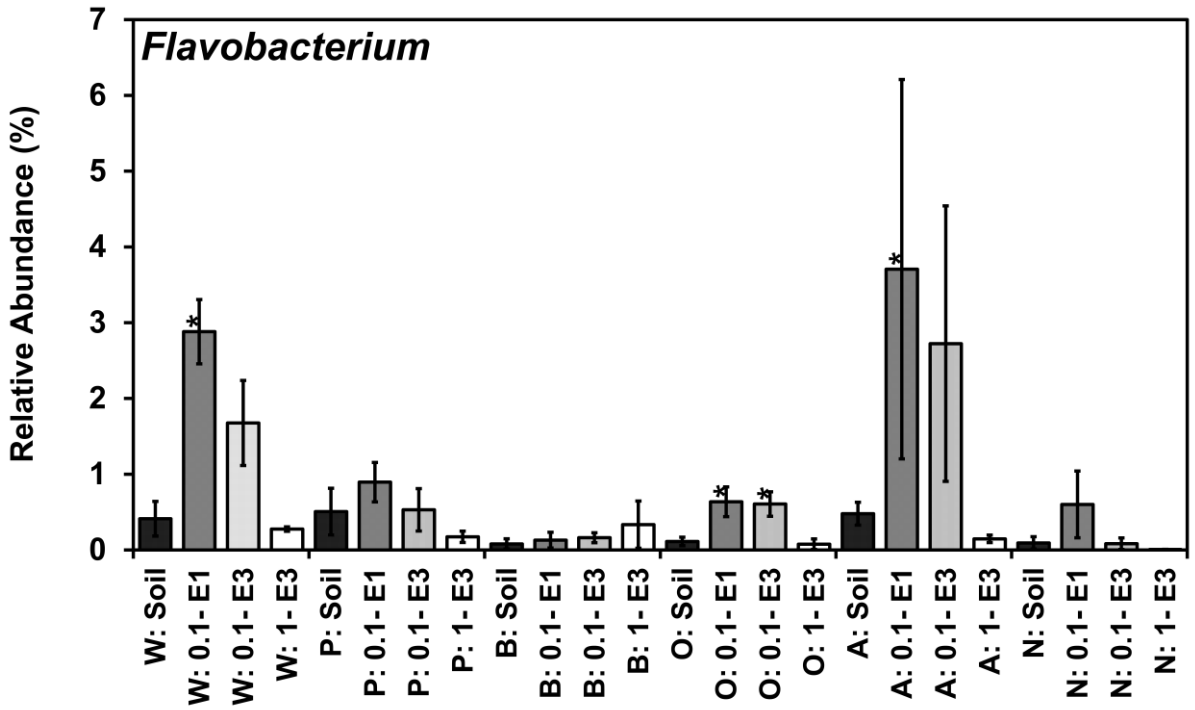


Figure 2.SI.25 *Flavobacterium* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 1.69 P=0.12), was affected by concentration ($F_{8,62}$ = 2.14 P=0.05), and was not affected by enrichment level ($F_{8,62}$ = 0.93 P=0.51), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Flavobacterium* defined by RDP classification at deepest assignment.

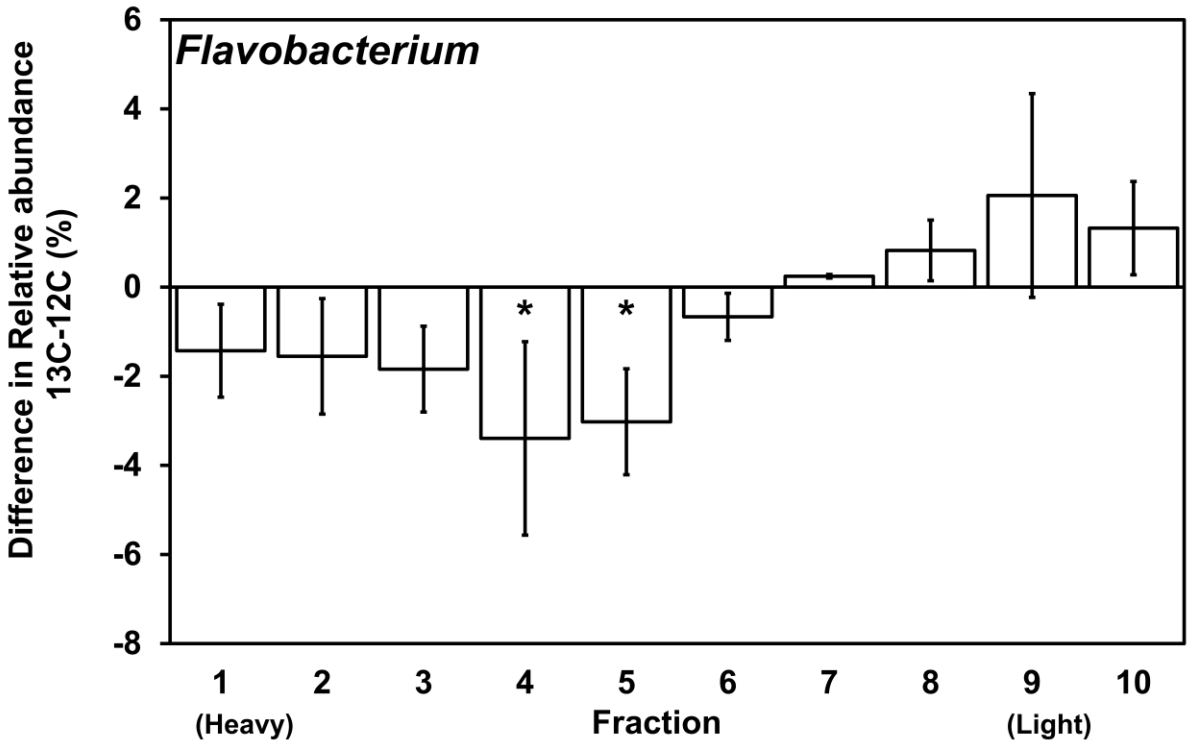


Figure 2.SI.26 *Flavobacterium* relative abundances in ^{13}C fractions net of *Flavobacterium* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Flavobacterium* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not significantly affected by the density ($F_{3,65}$ = 0.66 P=0.59), and was not affected by the carbon type ($F_{3,65}$ = 0.45 P=0.72), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

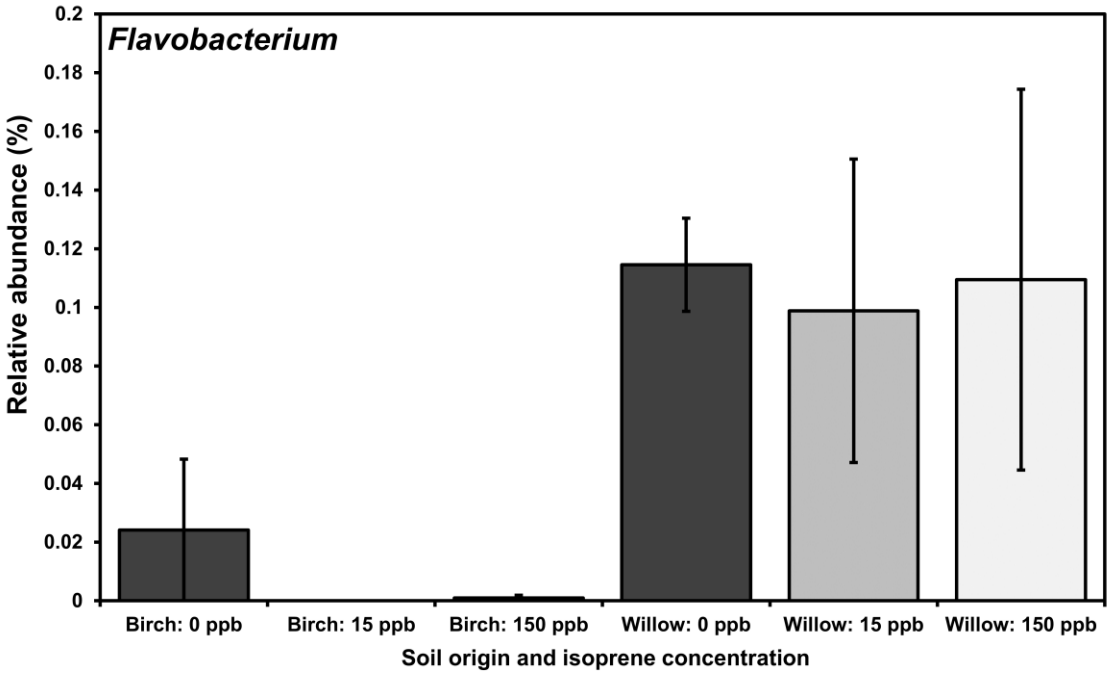


Figure 2.SI.27 *Flavobacterium* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Flavobacterium* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 0.56 P=0.8) and was not affected by the tree type ($F_{8,15}$ = 2.21 P=0.09) .

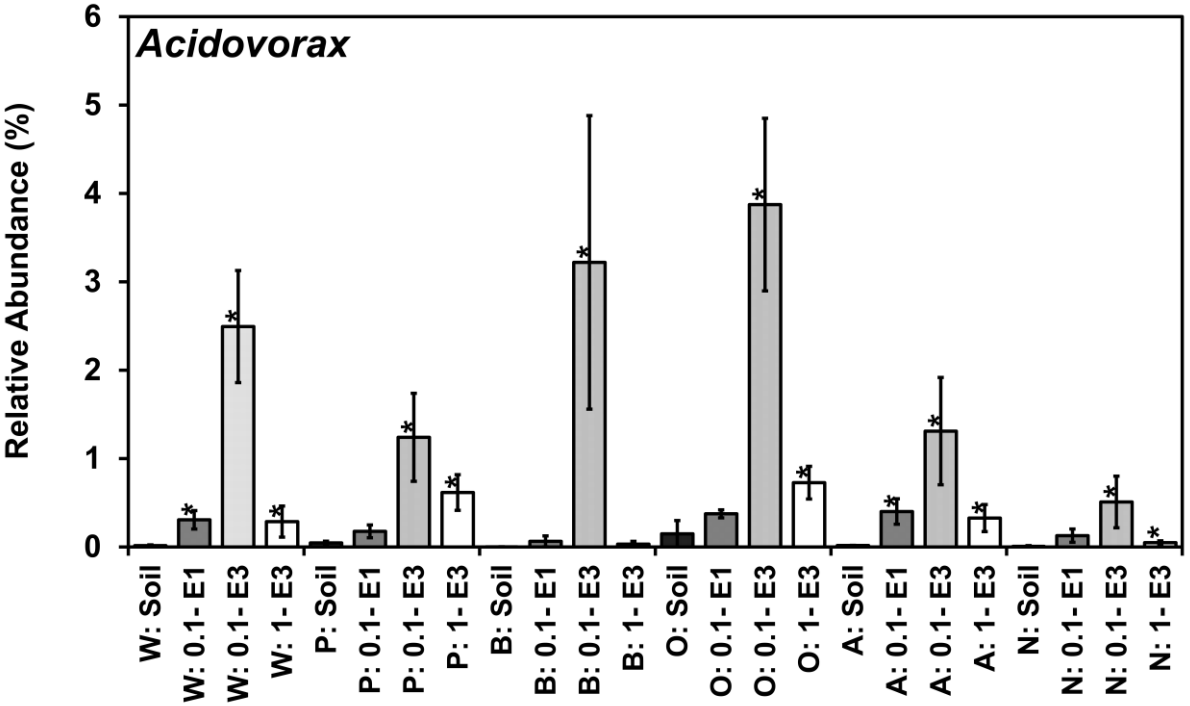


Figure 2.SI.28 *Acidovorax* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 0.99 P=0.46), was affected by concentration ($F_{8,62}$ = 9.26 P= < 0.001), and was affected by enrichment level ($F_{8,62}$ = 4.14 P= < 0.001), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Acidovorax* defined by RDP classification at deepest assignment.

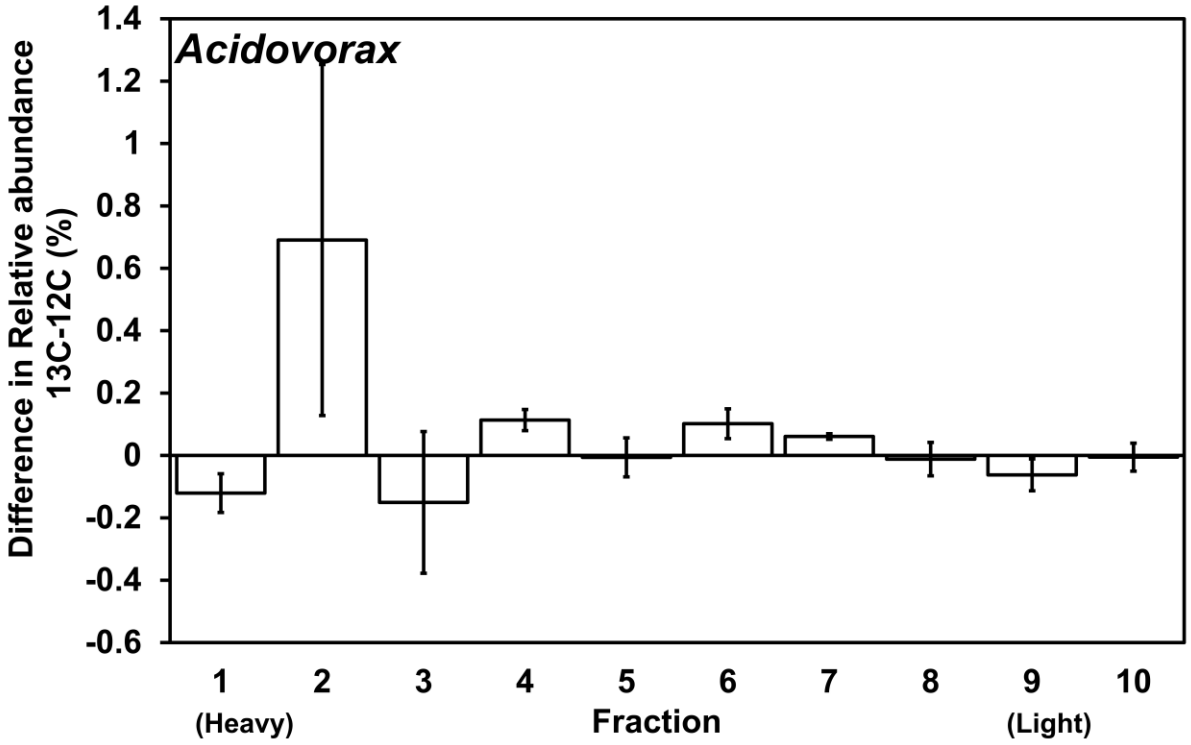


Figure 2.SI.29 *Acidovorax* relative abundances in ^{13}C fractions net of *Acidovorax* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Acidovorax* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was significantly affected by the density ($F_{3,65}$ = 3.7 P=0.02), and was not affected by the carbon type ($F_{3,65}$ = 0.59 P=0.63), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

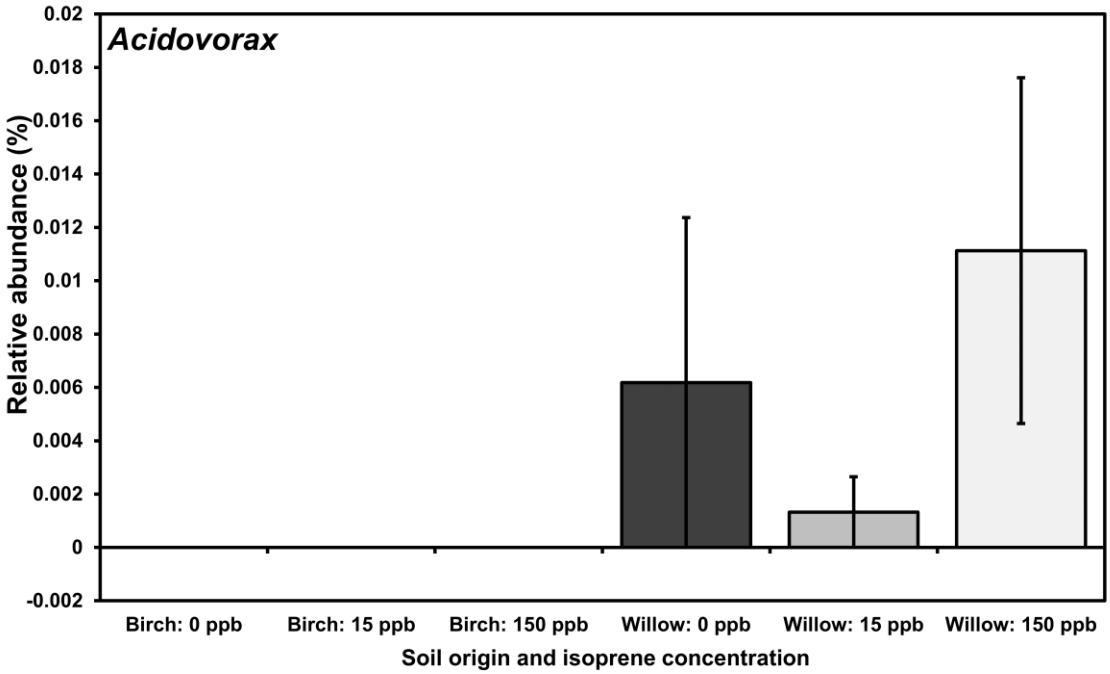


Figure 2.SI.30 *Acidovorax* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Acidovorax* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 0.01 *P*=1) and was not affected by the tree type ($F_{8,15}$ = 1.37 *P*=0.29).

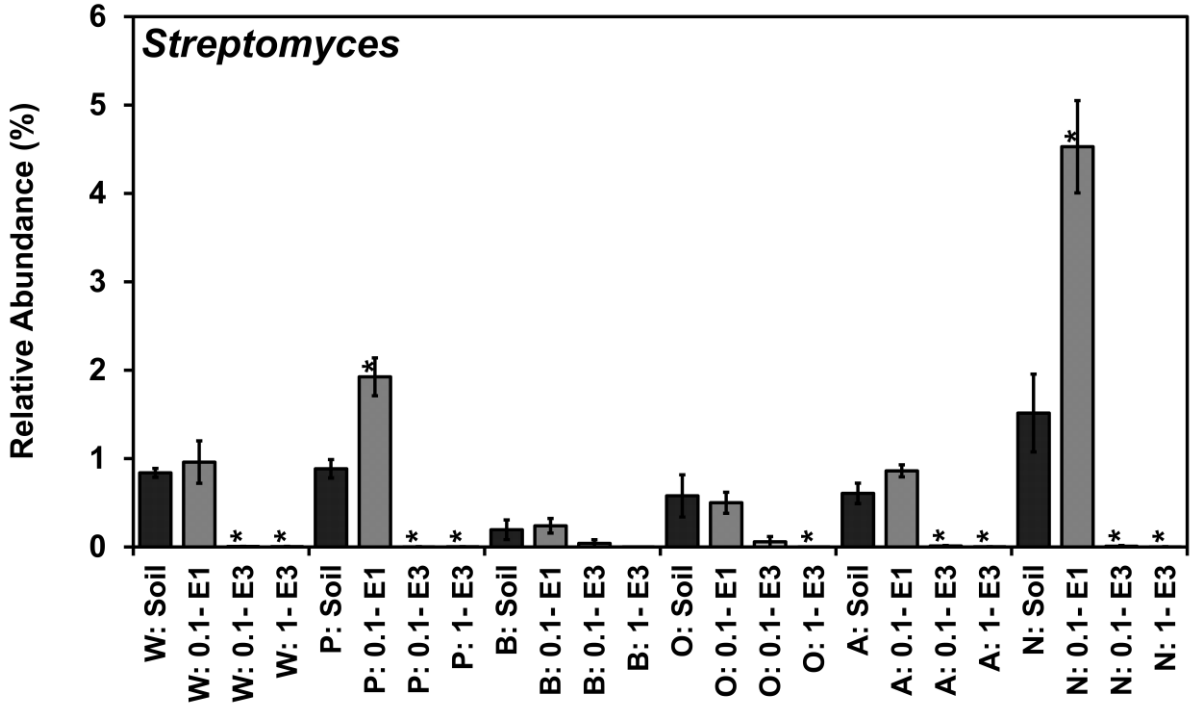


Figure 2.SI.31 *Streptomyces* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 1.04 P=0.42), not affected by concentration ($F_{8,62}$ = 0.52 P=0.84), and was affected by enrichment level ($F_{8,62}$ = 3.76 P= < 0.001), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Streptomyces* defined by RDP classification at deepest assignment.

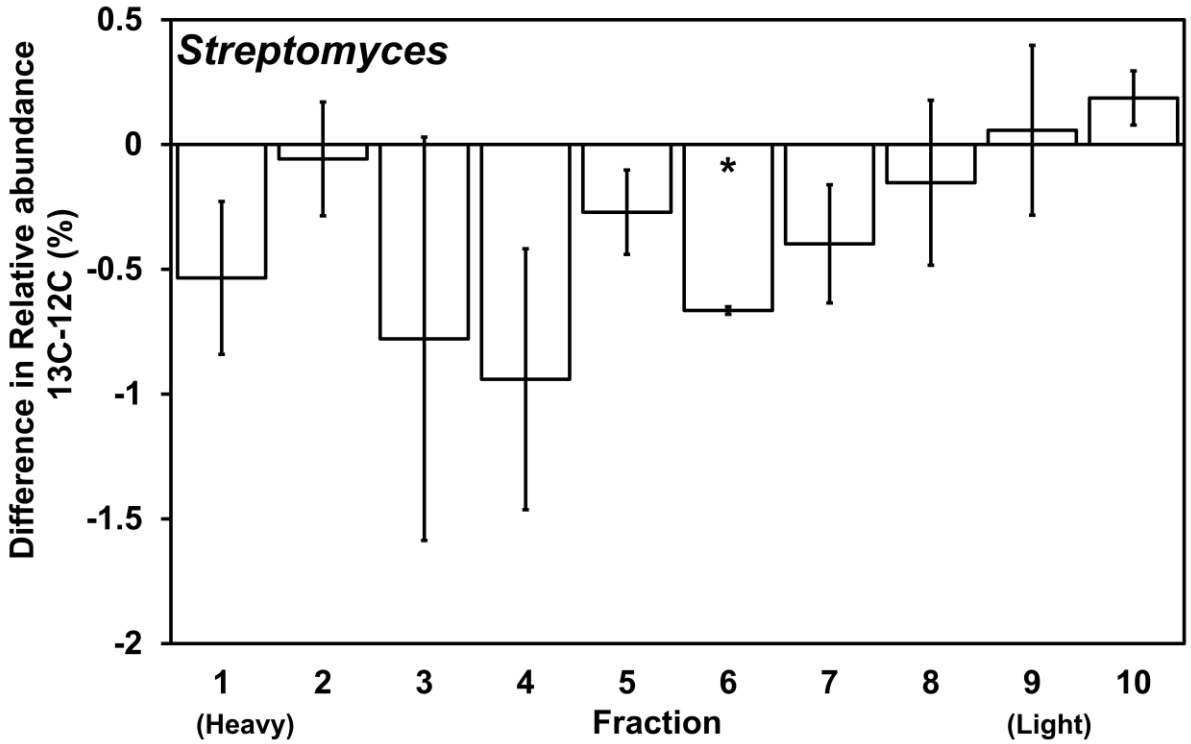


Figure 2.SI.32 *Streptomyces* relative abundances in ^{13}C fractions net of *Streptomyces* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Streptomyces* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not significantly affected by the density ($F_{3,65}$ = 1.1 P=0.36), and was not affected by the carbon type ($F_{3,65}$ = 0.71 P=0.55), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

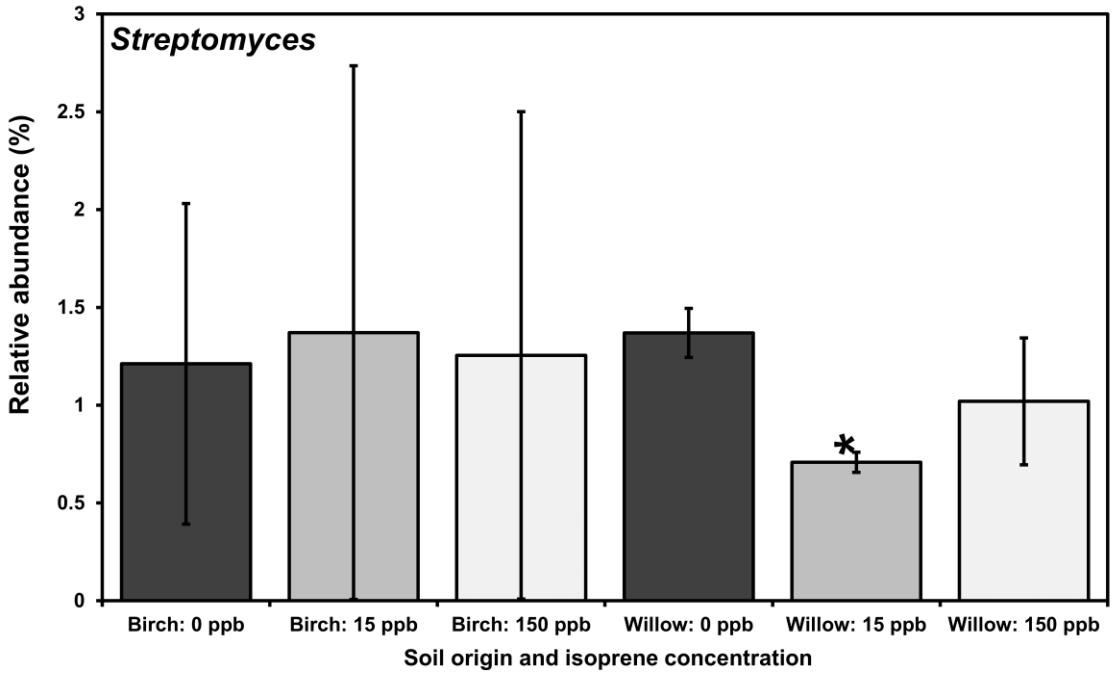


Figure 2.SI.33 *Streptomyces* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 1.27 P=0.33) and was not affected by the tree type ($F_{8,15}$ = 1.13 P=0.4).

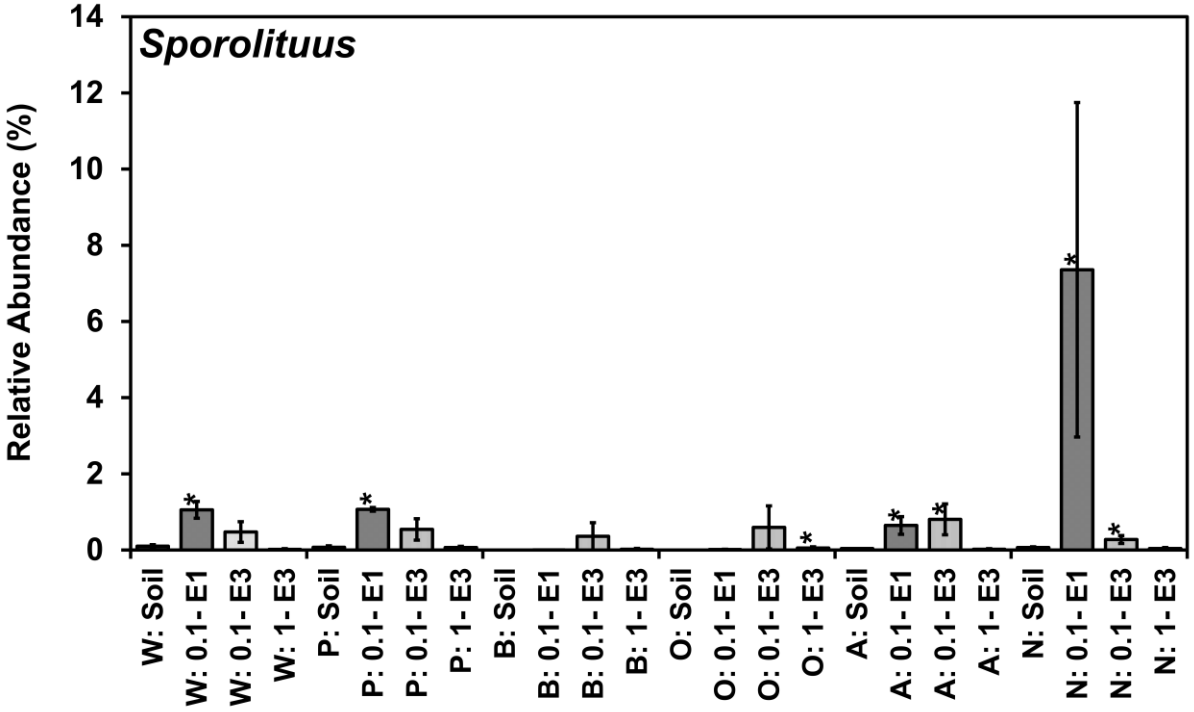


Figure 2.SI.34 *Sporolituus* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 0.72 P=0.68), not affected by concentration ($F_{8,62}$ = 1.03 P=0.43), and was not affected by enrichment level ($F_{8,62}$ = 0.33 P=0.95), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Sporolituus* defined by RDP classification at deepest assignment.

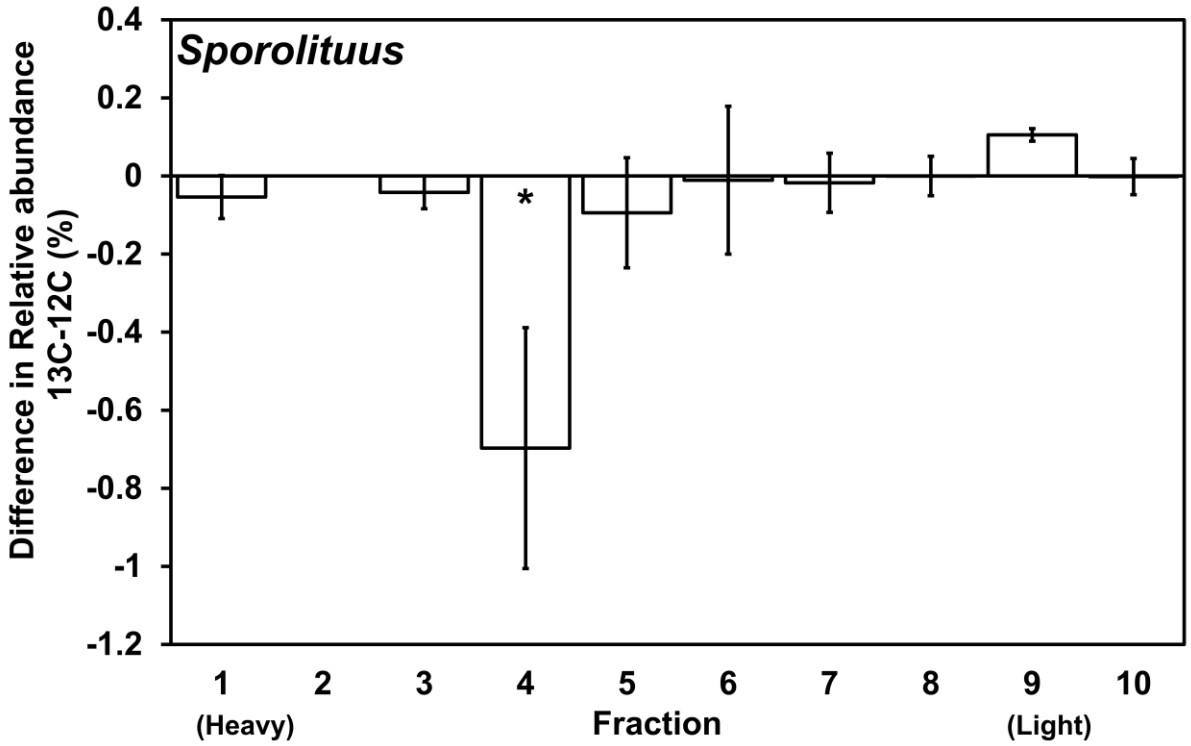


Figure 2.SI.35 *Sporolituus* relative abundances in ^{13}C fractions net of *Sporolituus* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Sporolituus* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was affected by the density ($F_{3,65}$ = 3.65 P=0.02), and was not affected by the carbon type ($F_{3,65}$ = 1.02 P=0.39), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

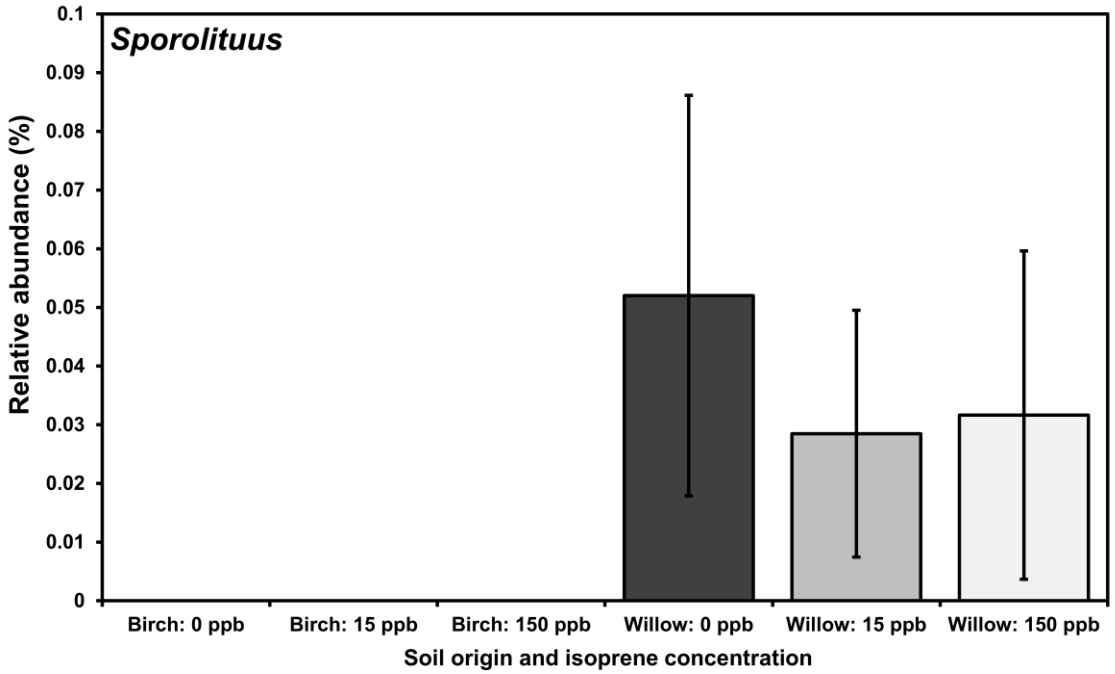


Figure 2.SI.36 *Sporolituus* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Sporolituus* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was affected by the isoprene concentration ($F_{8,15}$ = 0.01 *P*=1) and was not affected by the tree type ($F_{8,15}$ = 1.63 *P*=0.2).

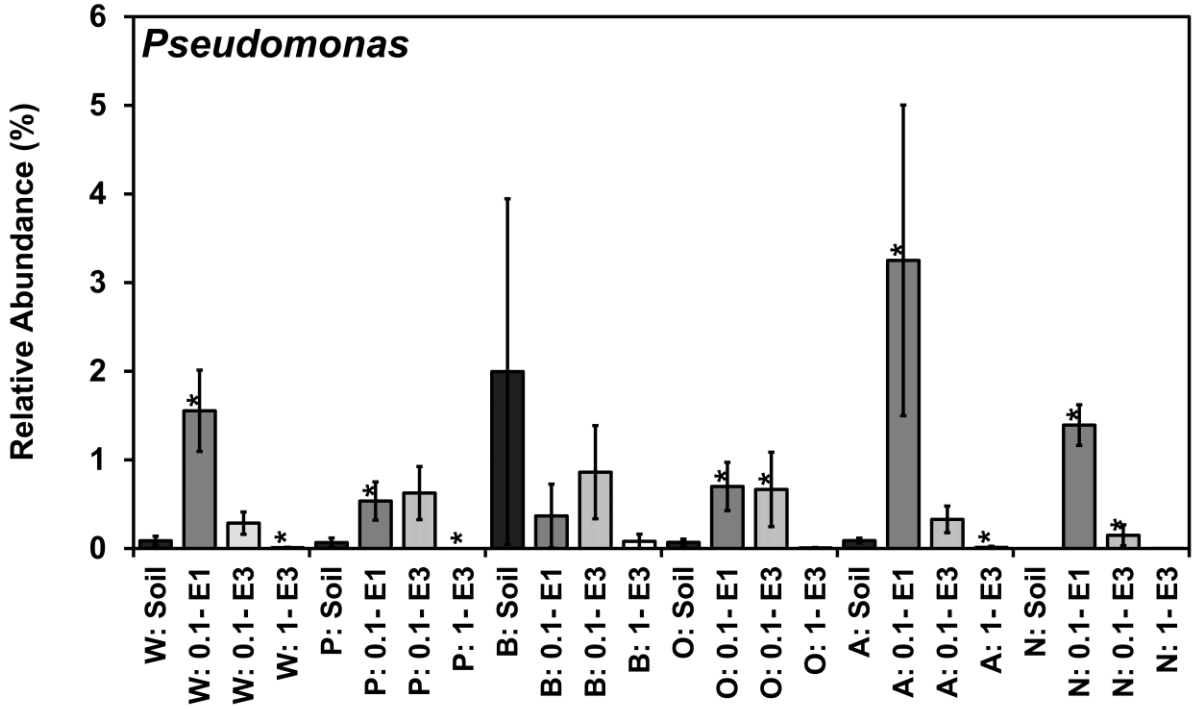


Figure 2.SI.37 *Pseudomonas* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 0.81 P=0.6), not affected by concentration ($F_{8,62}$ = 1.92 P=0.07), and was not affected by enrichment level ($F_{8,62}$ = 1.73 P=0.11), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Pseudomonas* defined by RDP classification at deepest assignment.

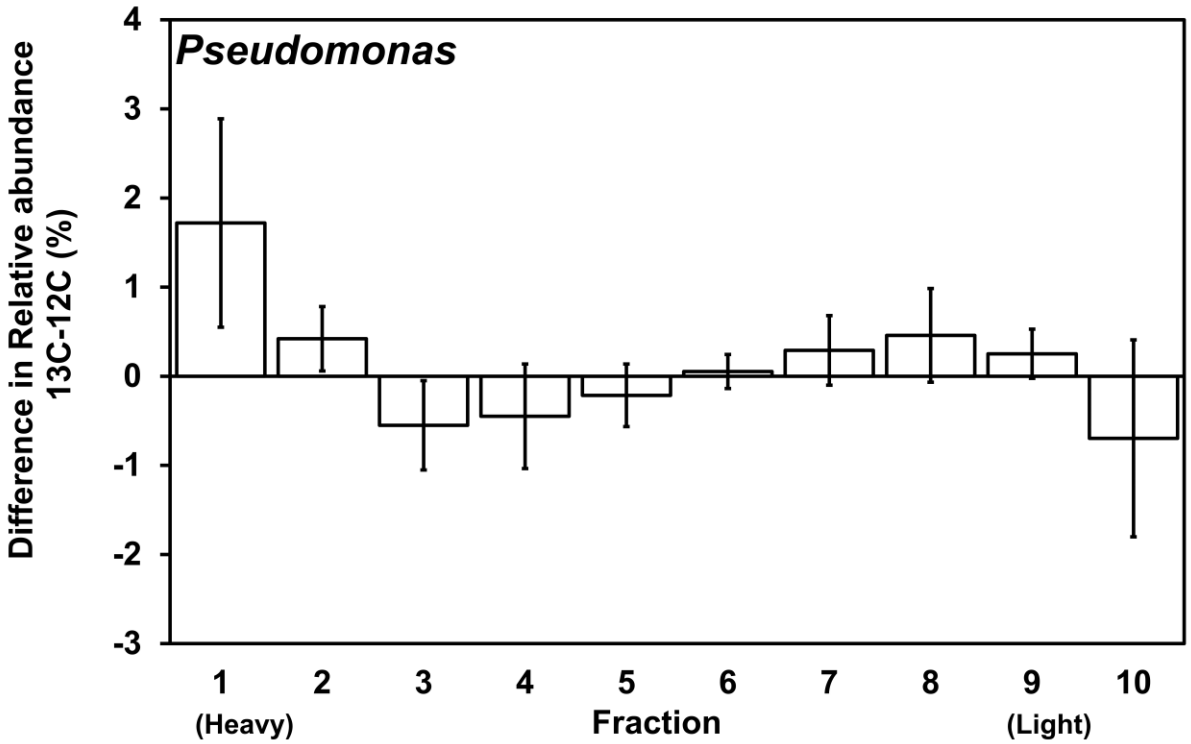


Figure 2.SI.38 *Pseudomonas* relative abundances in ^{13}C fractions net of *Pseudomonas* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Pseudomonas* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not significantly affected by the density ($F_{3,65}$ = 0.4 P=0.76), and was not affected by the carbon type ($F_{3,65}$ = 0.81 P=0.49), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

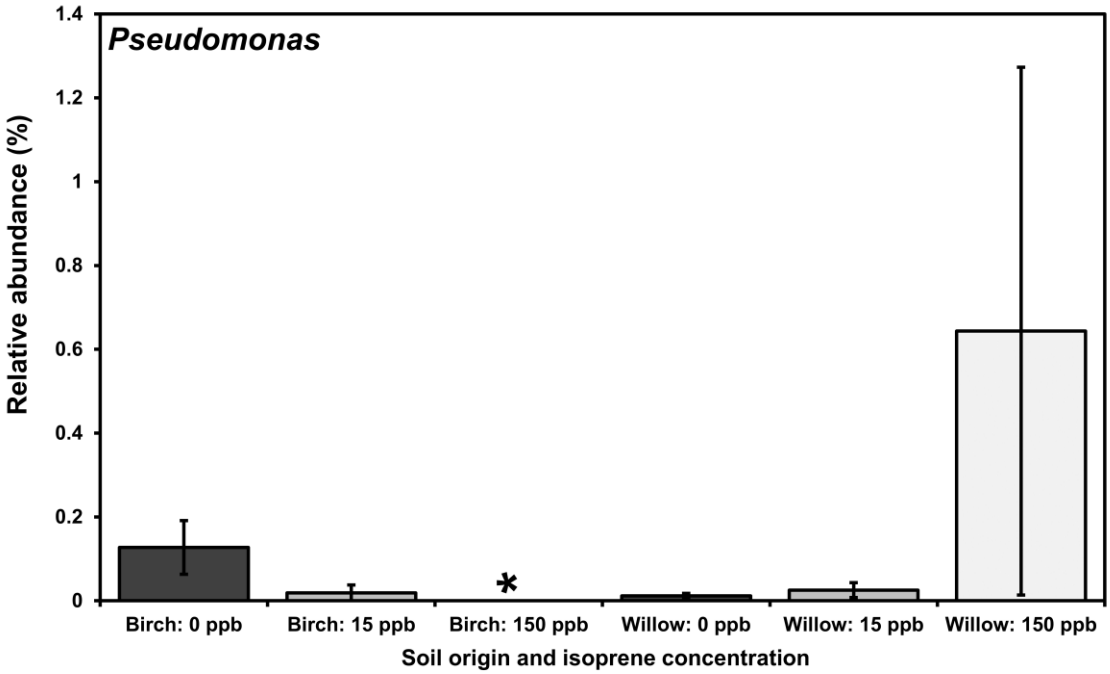


Figure 2.SI.39 *Pseudomonas* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Pseudomonas* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 0.5 P=0.84) and was not affected by the tree type ($F_{8,15}$ = 0.53 P=0.82).

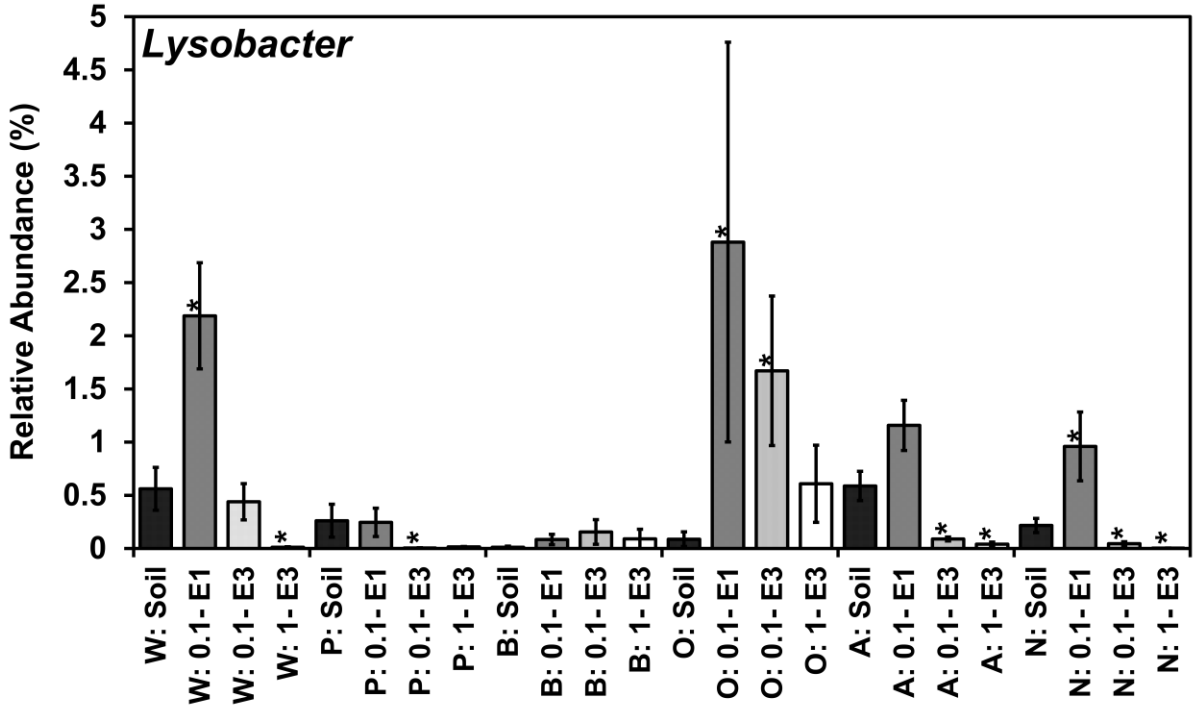


Figure 2.SI.40 *Lysobacter* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3.A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 1.13 P=0.36), not affected by concentration ($F_{8,62}$ = 1.39 P=0.22), and was not affected by enrichment level ($F_{8,62}$ = 1.14 P=0.35), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Lysobacter* defined by RDP classification at deepest assignment.

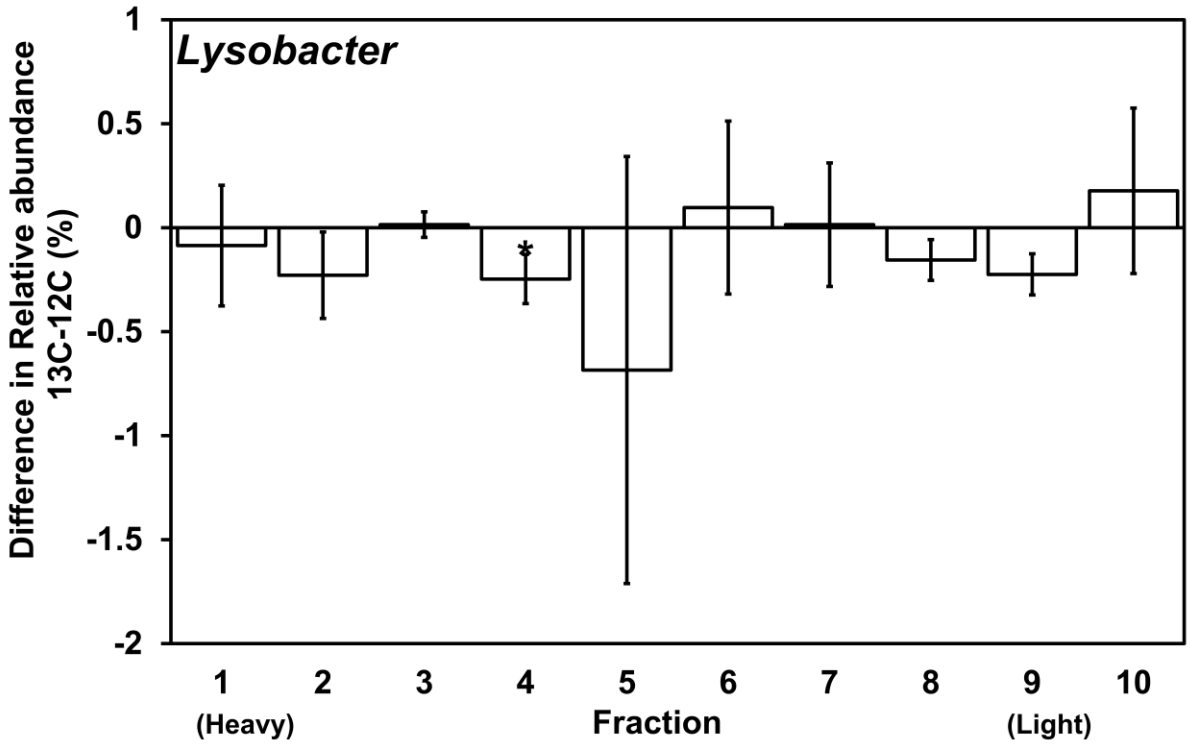


Figure 2.SI.41 *Lysobacter* relative abundances in ^{13}C fractions net of *Lysobacter* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Lysobacter* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not significantly affected by the density ($F_{3,65}$ = 1.68 P=0.18), and was not affected by the carbon type ($F_{3,65}$ = 0.85 P=0.47), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

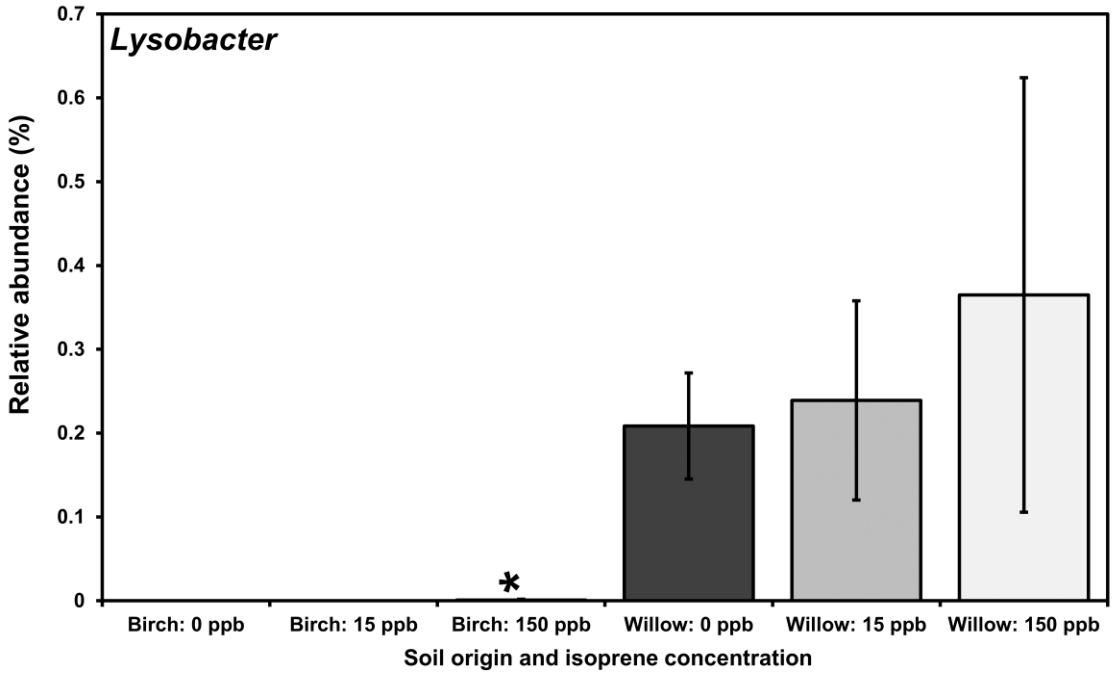


Figure 2.SI.42 *Lysobacter* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Lysobacter* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 0.11 P=1) and was not affected by the tree type ($F_{8,15}$ = 1.42 P=0.27).

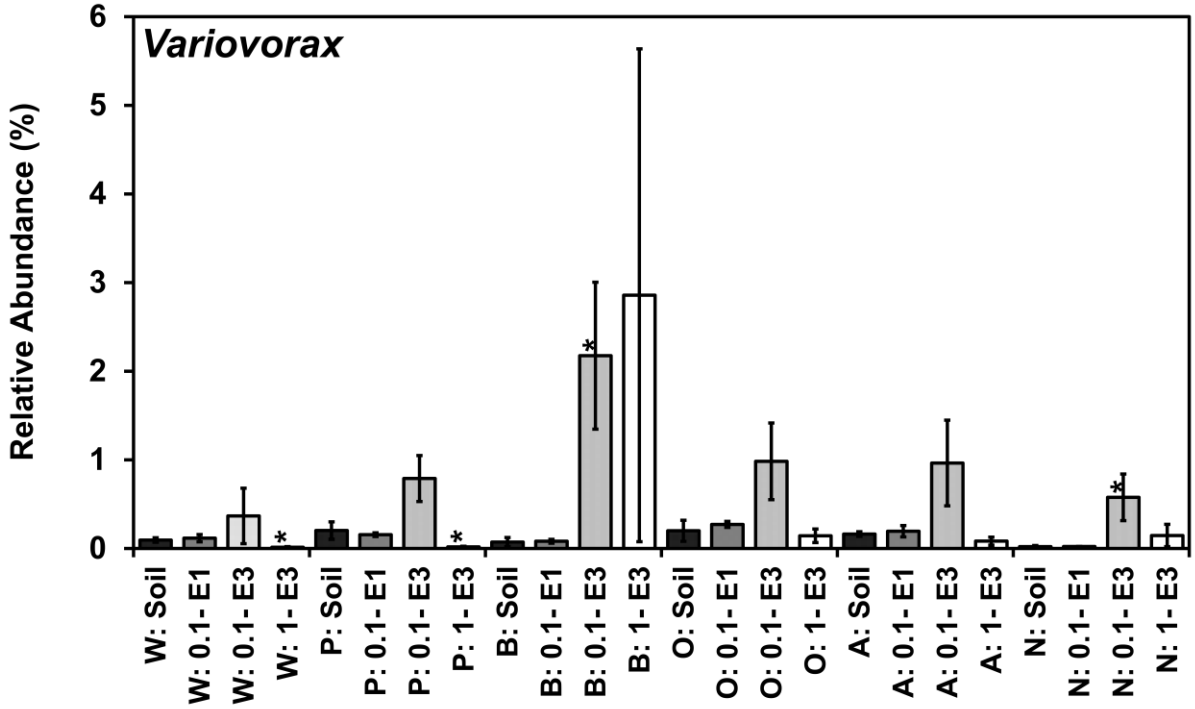


Figure 2.SI.43 *Variovorax* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 0.77 P=0.63), not affected by concentration ($F_{8,62}$ = 1.17 P=0.33), and was not affected by enrichment level ($F_{8,62}$ = 1.17 P=0.33), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Variovorax* defined by RDP classification at deepest assignment.

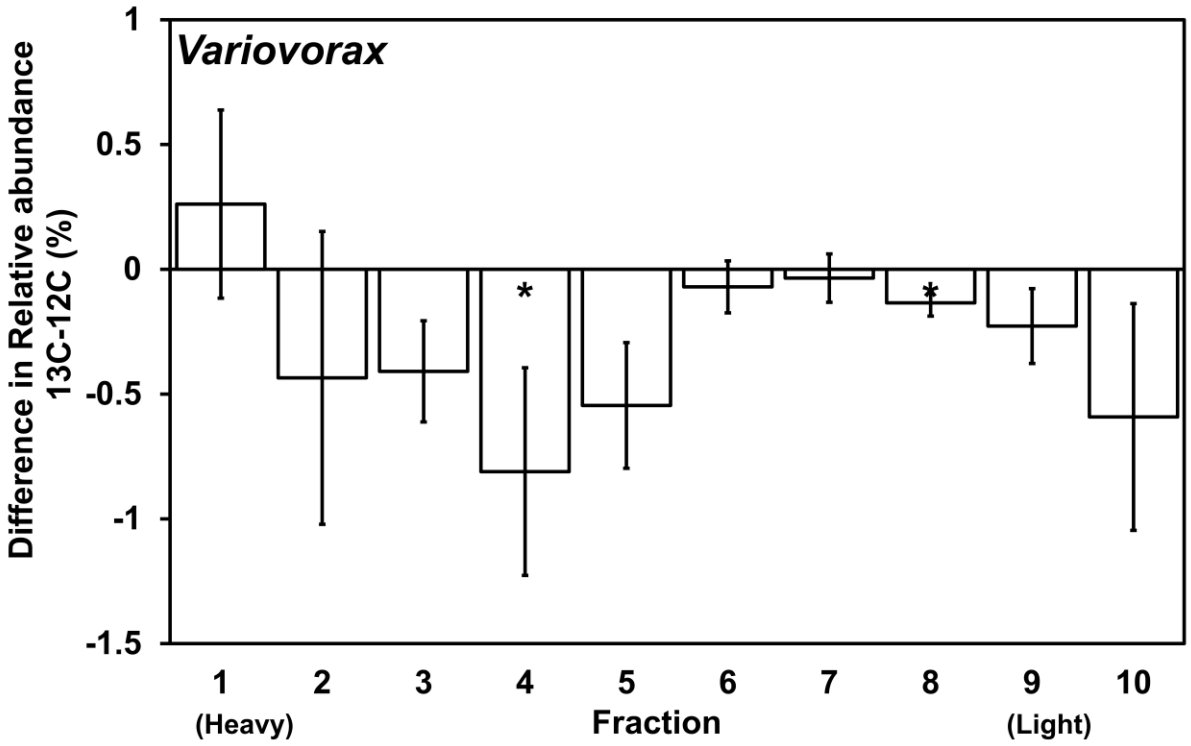


Figure 2.SI.44 *Variovorax* relative abundances in ^{13}C fractions net of *Variovorax* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Variovorax* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not significantly affected by the density ($F_{3,65}$ = 0.86 P=0.47), and was not affected by the carbon type ($F_{3,65}$ = 0.82 P=0.49), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

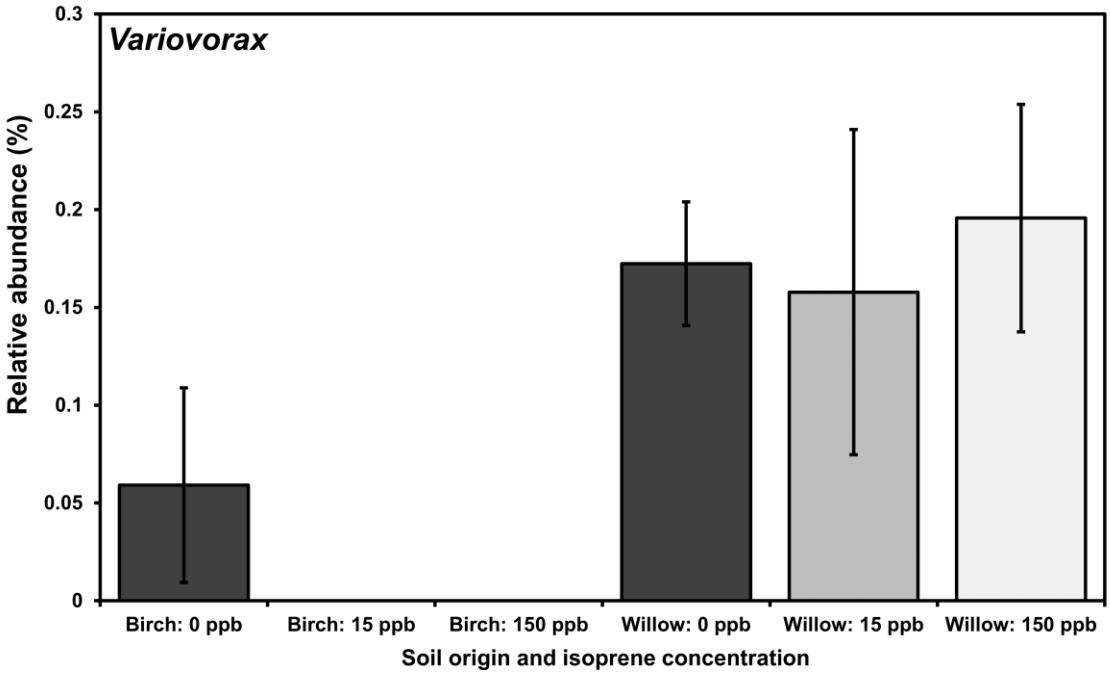


Figure 2.SI.45 *Variovorax* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Variovorax* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 1.08 P=0.43) and was not affected by the tree type ($F_{8,15}$ = 1.41 P=0.27).

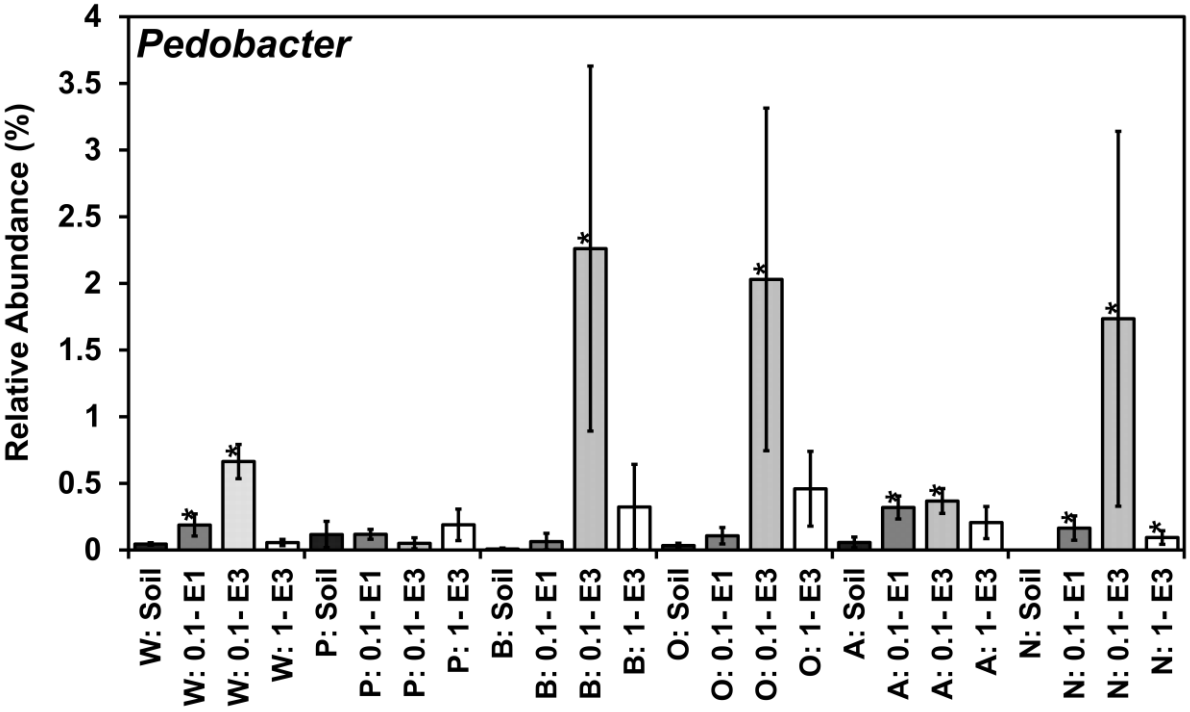


Figure 2.SI.46 *Pedobacter* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 0.91 P=0.52), was affected by concentration ($F_{8,62}$ = 3.55 P= < 0.001), and was affected by enrichment level ($F_{8,62}$ = 2.12 P=0.05), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Pedobacter* defined by RDP classification at deepest assignment.

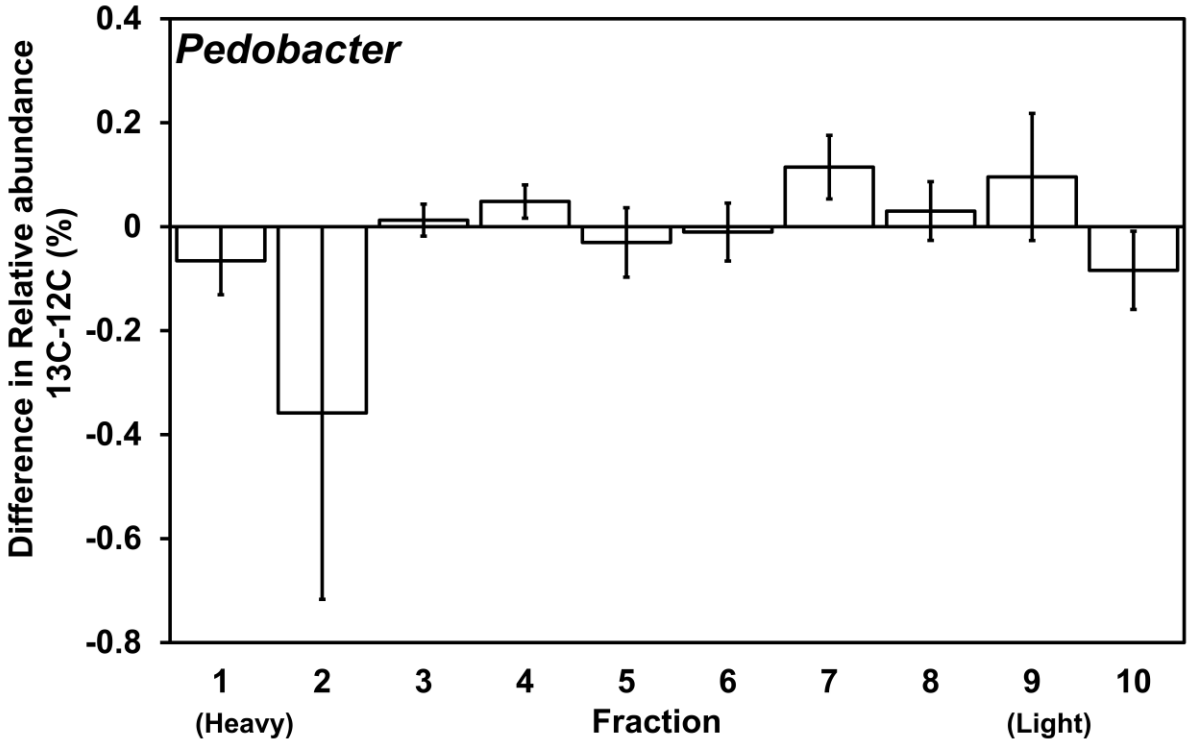


Figure 2.SI.47 *Pedobacter* relative abundances in ^{13}C fractions net of *Pedobacter* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Pedobacter* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not significantly affected by the density ($F_{3,65}$ = 0.29 P=0.84), and was not affected by the carbon type ($F_{3,65}$ = 0.59 P=0.63), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

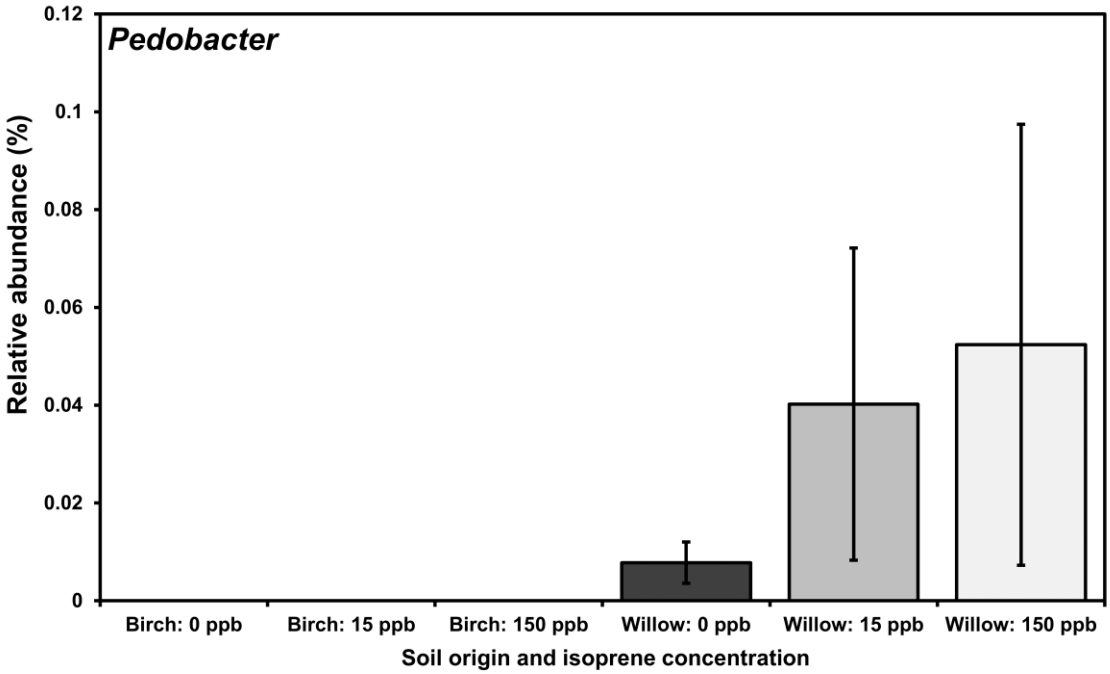


Figure 2.SI.48 *Pedobacter* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Pedobacter* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 0.01 P=1) and was not affected by the tree type ($F_{8,15}$ = 0.54 P=0.81).

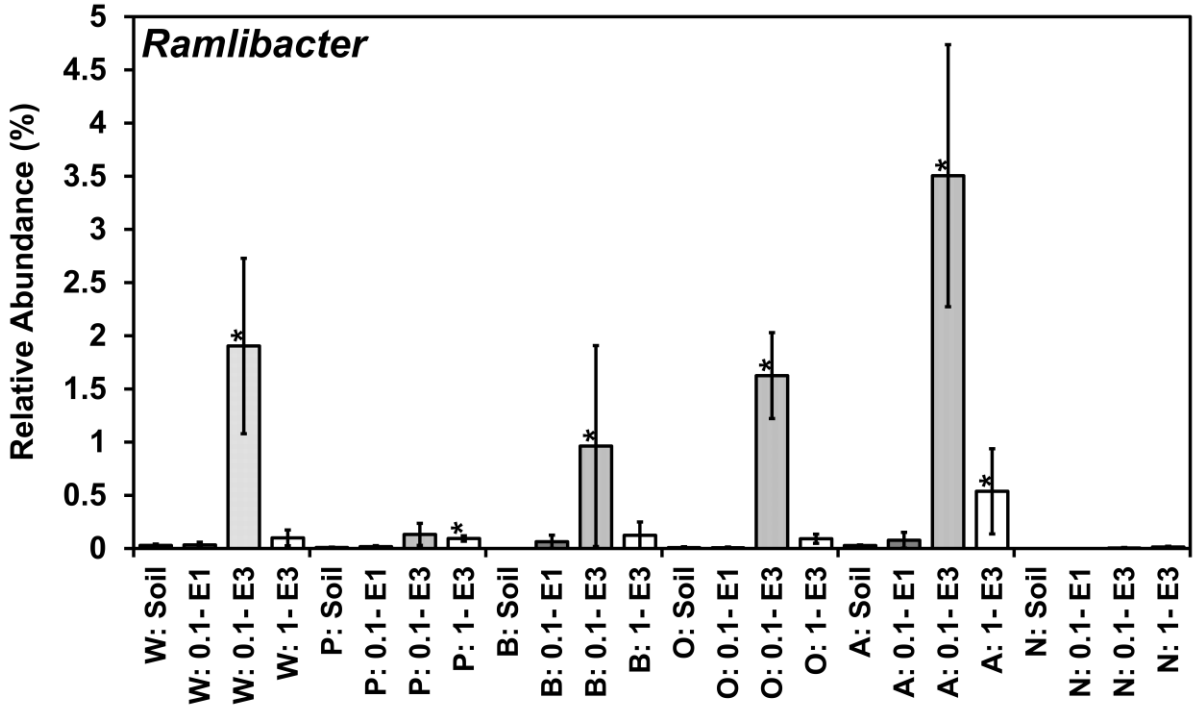


Figure 2.SI.49 *Ramlibacter* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3.A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was affected by sample location ($F_{8,62}$ = 3.19 P=0.04), was affected by concentration ($F_{8,62}$ = 5.46 P= < 0.001), and was not affected by enrichment level ($F_{8,62}$ = 1.25 P=0.29), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Ramlibacter* defined by RDP classification at deepest assignment.

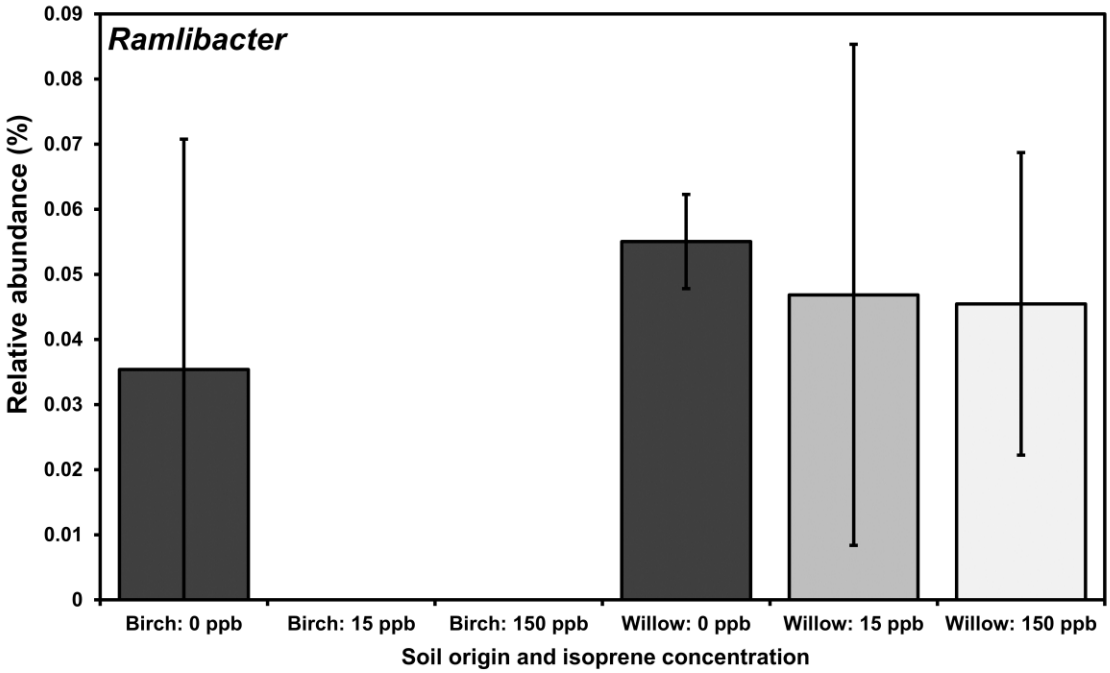


Figure 2.SI.51 *Ramlibacter* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Ramlibacter* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was was significantly affected by the density ($F_{3,65}$ = 0.41 *P*=0.36), and was was affected by the carbon type ($F_{3,65}$ = 0.57 *P*= < 0.001), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

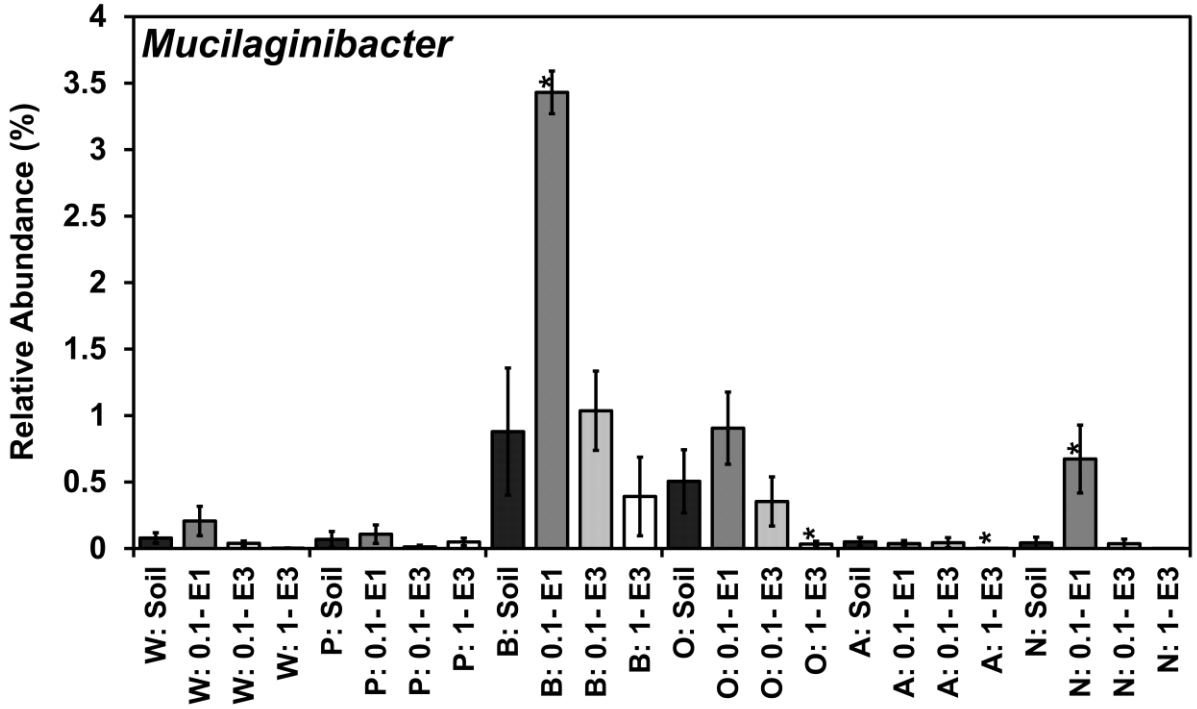


Figure 2.SI.52 *Mucilaginibacter* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 1.02 P=0.44), was affected by concentration ($F_{8,62}$ = 0.88 P=0.54), and was affected by enrichment level ($F_{8,62}$ = 3.19 P=0.03), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Mucilaginibacter* defined by RDP classification at deepest assignment.

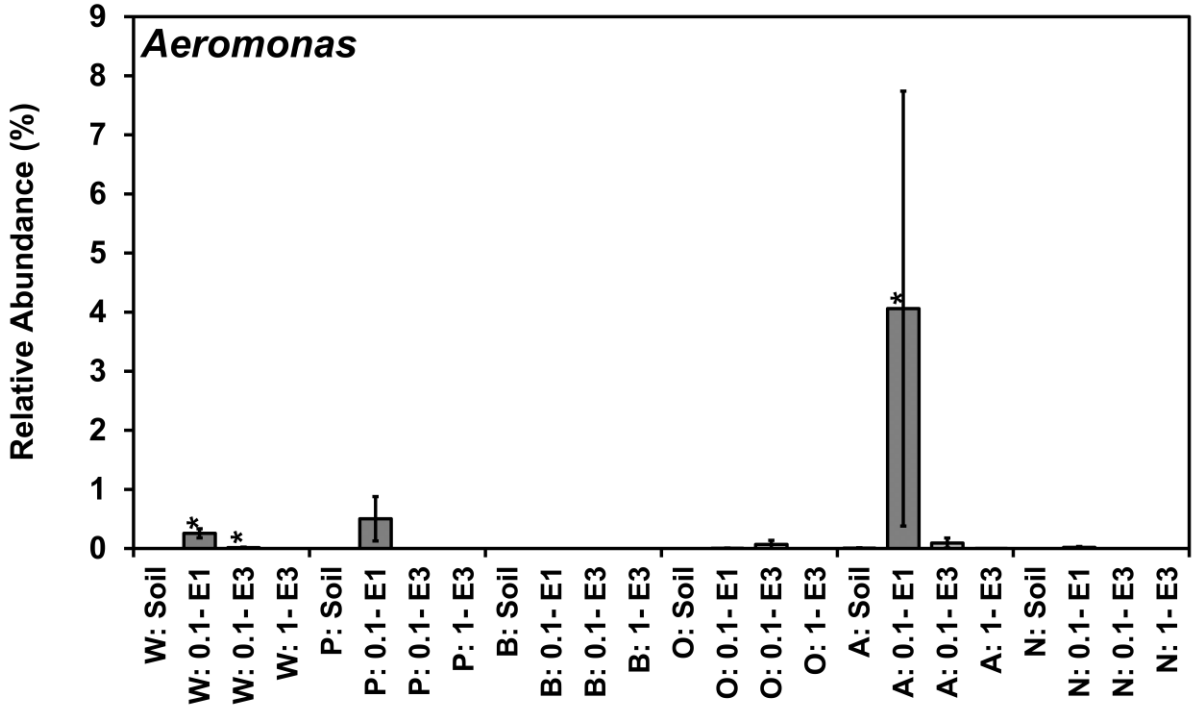


Figure 2.SI.53 *Aeromonas* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 1.92 P=0.07), not affected by concentration (), and was not affected by enrichment level ($F_{8,62}$ = 0.27 P=0.98), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Aeromonas* defined by RDP classification at deepest assignment.

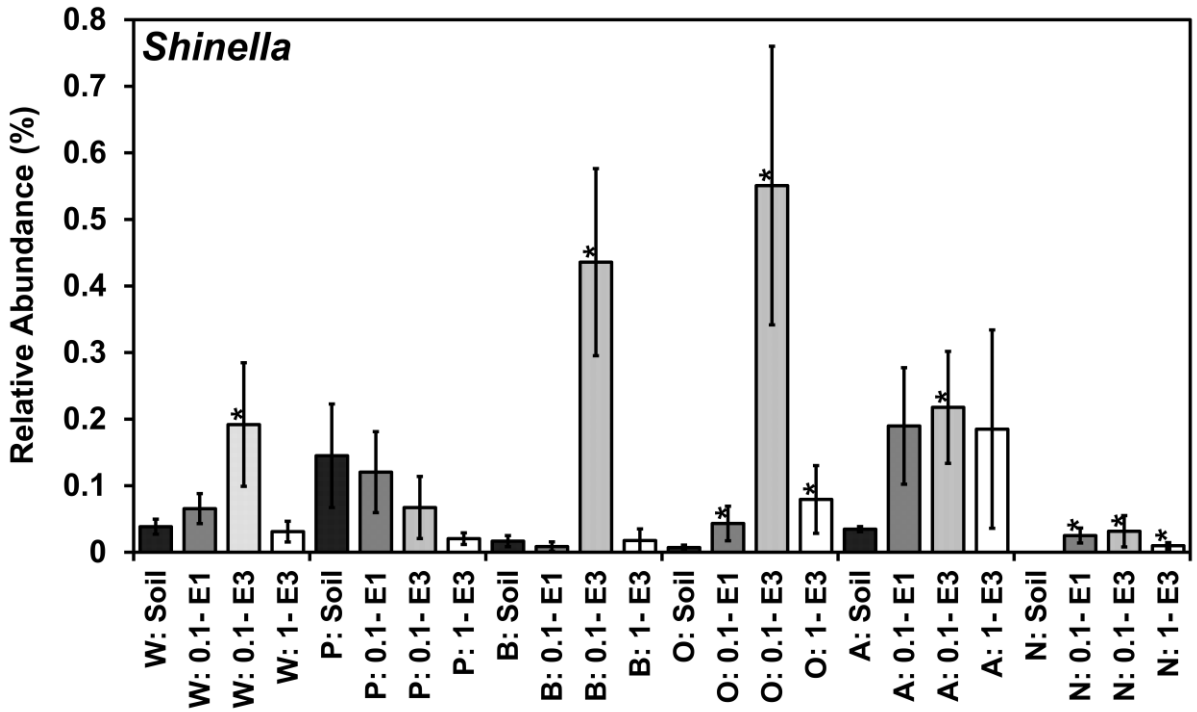


Figure 2.SI.54 *Shinella* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3.A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 1.08 P=0.39), was affected by concentration ($F_{8,62}$ = 4.23 P= < 0.001), and was not affected by enrichment level ($F_{8,62}$ = 1.14 P=0.35), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Shinella* defined by RDP classification at deepest assignment.

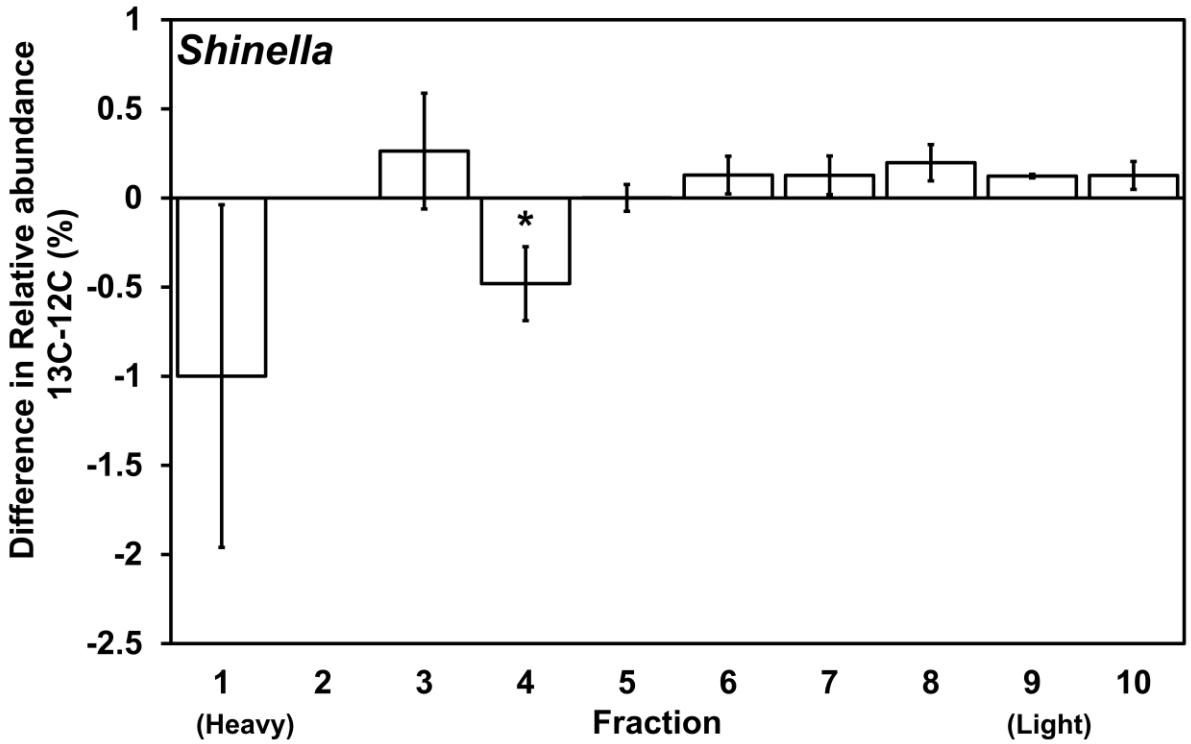


Figure 2.SI.55 *Shinella* relative abundances in ^{13}C fractions net of *Shinella* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Shinella* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not significantly affected by the density ($F_{3,65}$ = 1.09 P=0.36), and was not affected by the carbon type ($F_{3,65}$ = 0.57 P=0.64), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

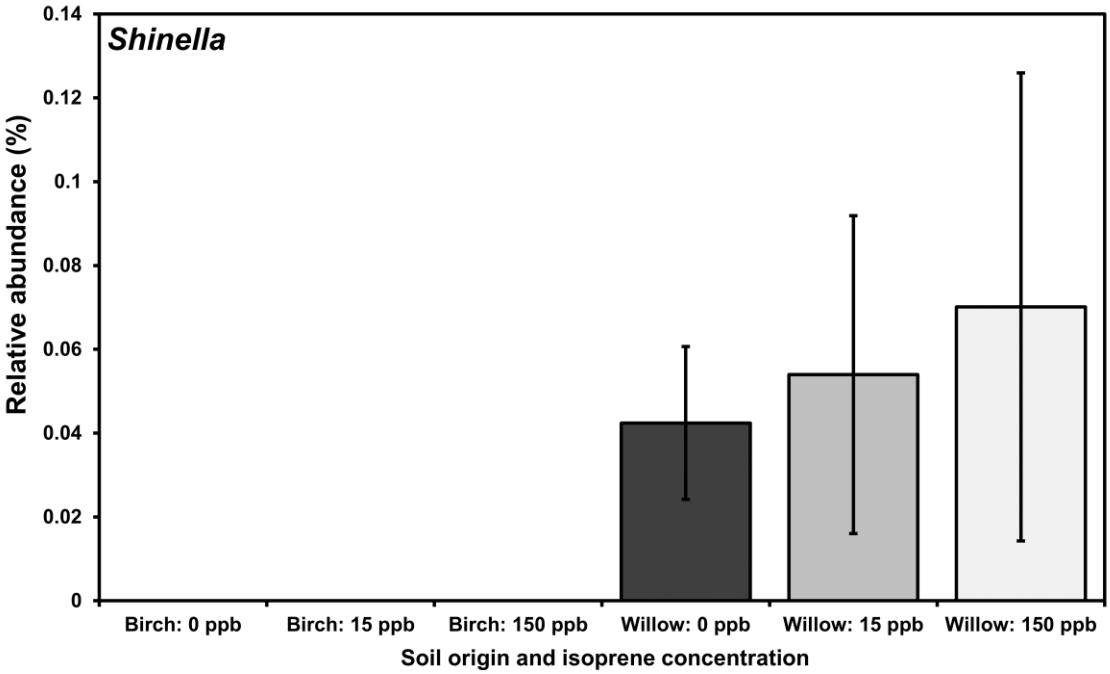


Figure 2.SI.56 *Shinella* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Shinella* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 0.01 P=1) and was not affected by the tree type ($F_{8,15}$ = 1.3 P=0.32).

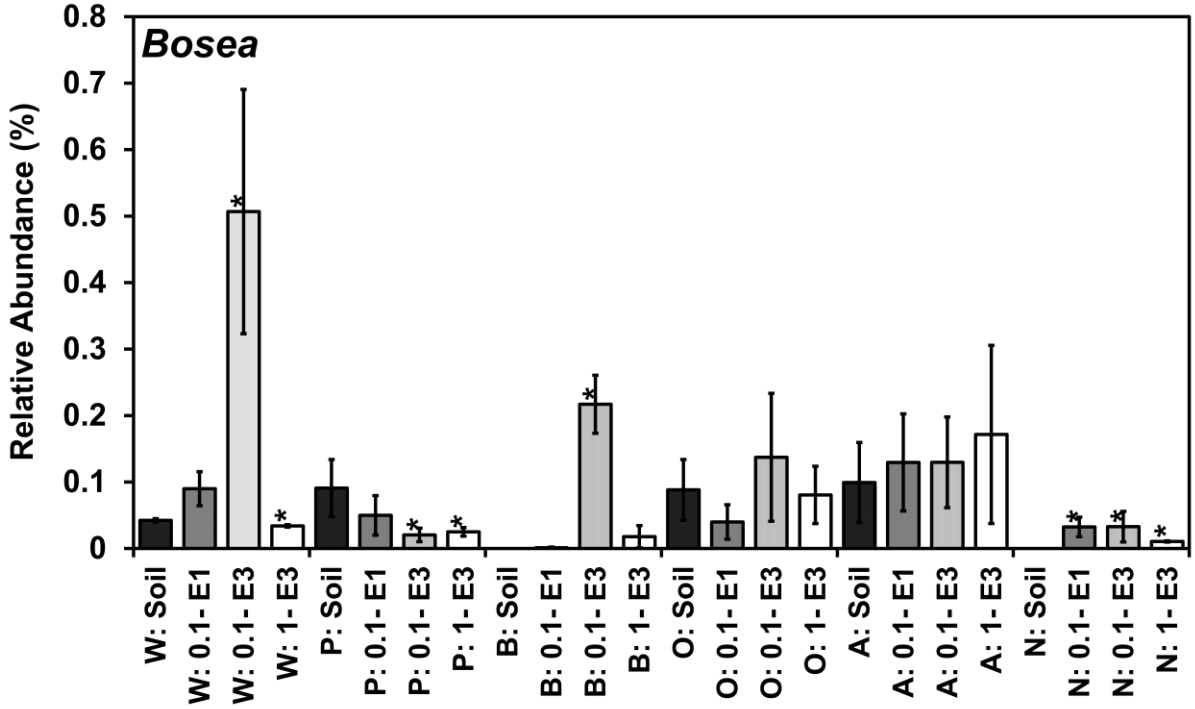


Figure 2.SI.57 *Bosea* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 1.38 P=0.22), was affected by concentration ($F_{8,62}$ = 3.19 P=0.01), and was not affected by enrichment level ($F_{8,62}$ = 2.03 P=0.06), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Bosea* defined by RDP classification at deepest assignment.

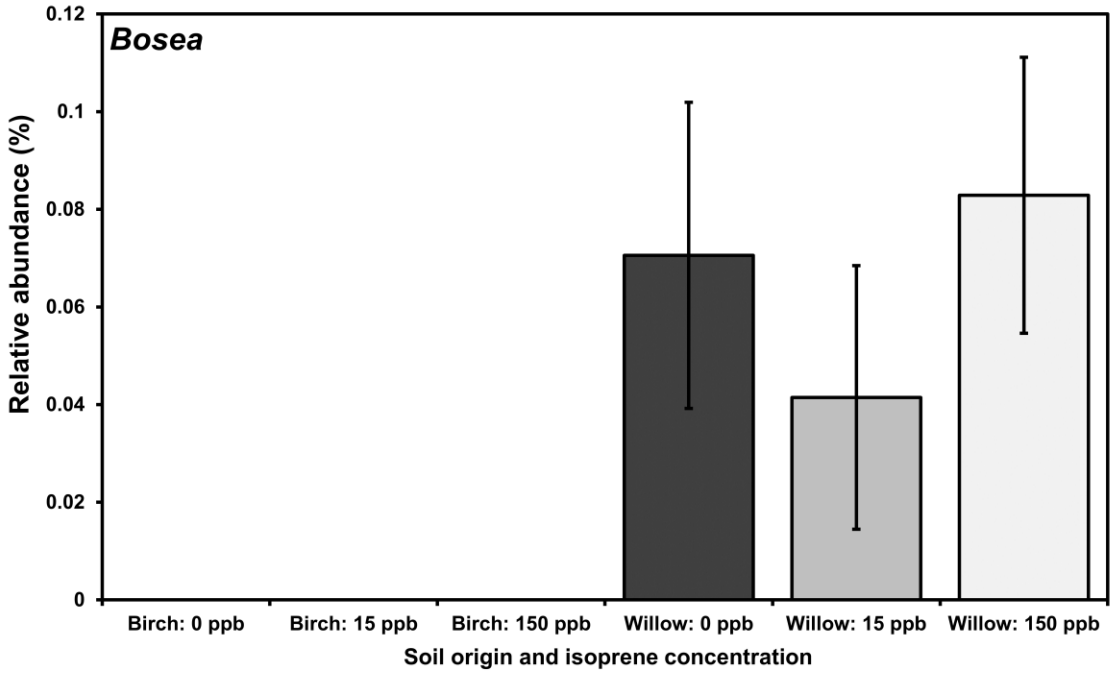


Figure 2.SI.58 *Bosea* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Bosea* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 0.01 P=1) and was not affected by the tree type ($F_{8,15}$ = 1.87 P=0.14).

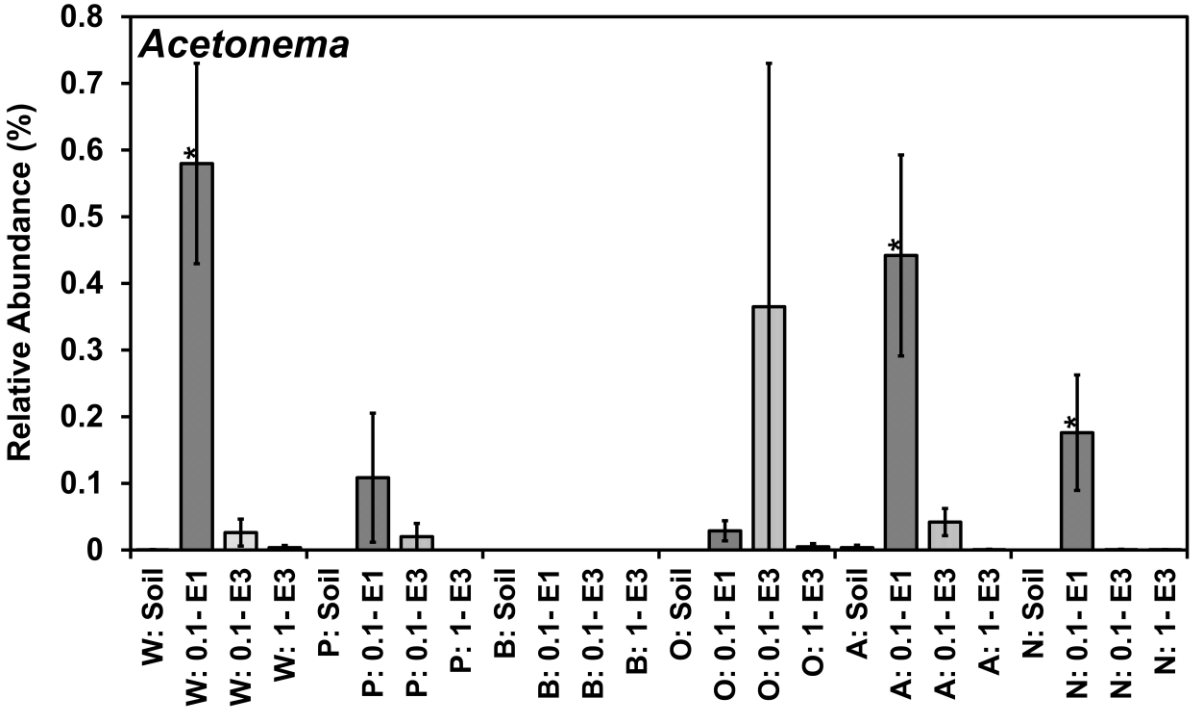


Figure 2.SI.59 *Acetonema* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3.A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 1.04 P=0.42), not affected by concentration ($F_{8,62}$ = 1.41 P=0.21), and was not affected by enrichment level ($F_{8,62}$ = 1.74 P=0.11), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Acetonema* defined by RDP classification at deepest assignment.

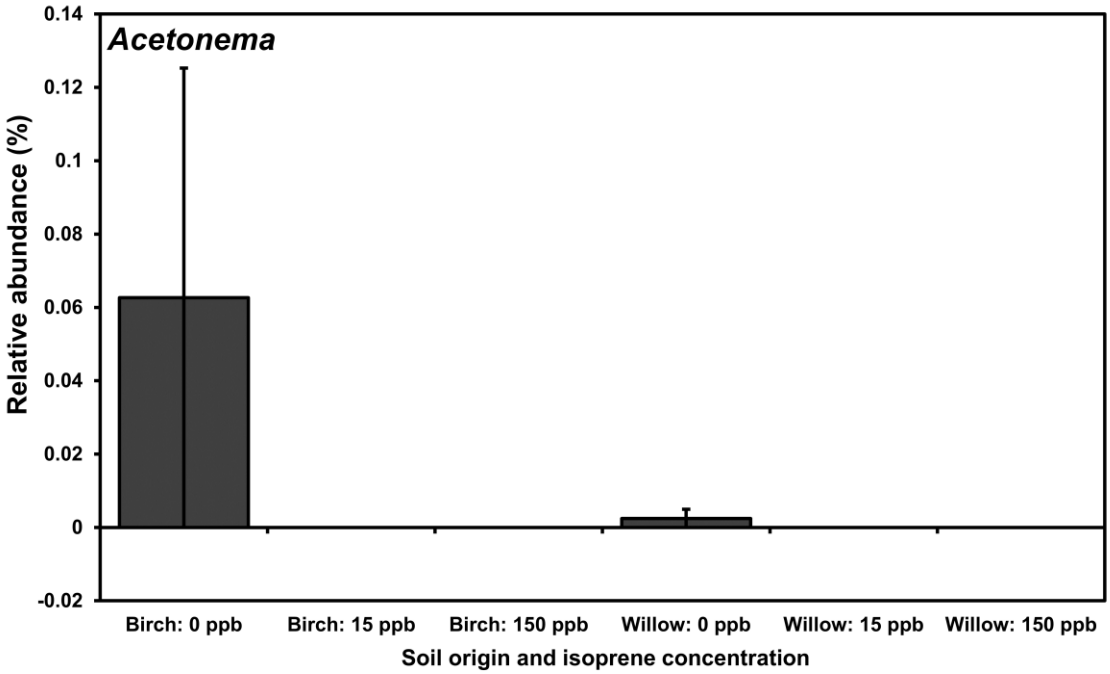


Figure 2.SI.60 *Acetonema* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Acetonema* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was affected by the isoprene concentration ($F_{8,15}$ = 3.39 P=0.02) and was affected by the tree type ($F_{8,15}$ = 3.94 P=0.01).

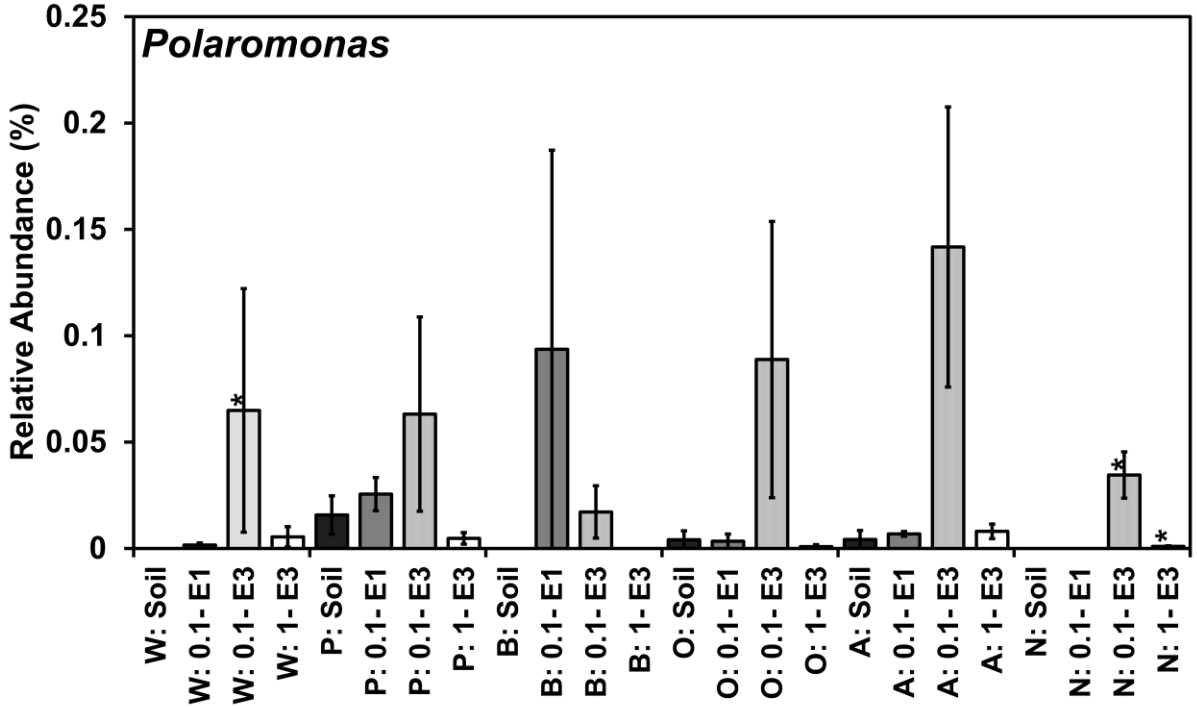


Figure 2.SI.61 *Polaromonas* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 1.04 P=0.42), was affected by concentration ($F_{8,62}$ = 3.95 P= < 0.001), and was not affected by enrichment level ($F_{8,62}$ = 1.92 P=0.07), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Polaromonas* defined by RDP classification at deepest assignment.

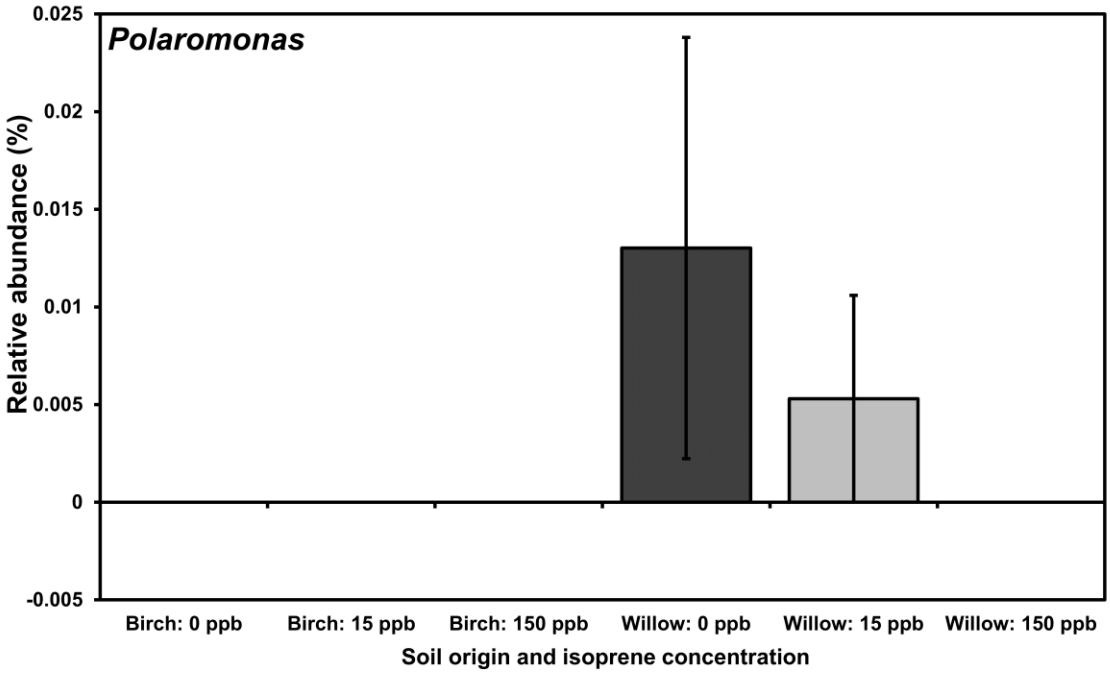


Figure 2.SI.62 *Polaromonas* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Polaromonas* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 0.01 P=1) and was not affected by the tree type ($F_{8,15}$ = 2.25 P=0.08).

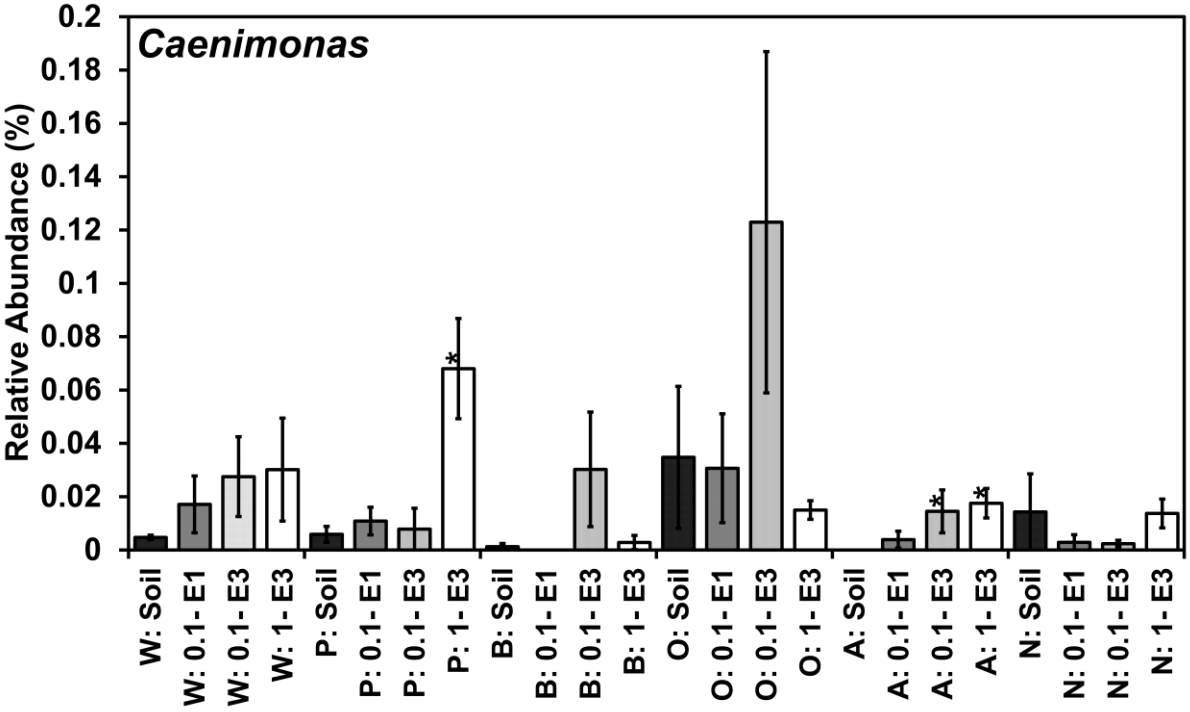


Figure 2.SI.63 *Caenimonas* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 0.9 P=0.53), not affected by concentration ($F_{8,62}$ = 1.02 P=0.44), and was not affected by enrichment level ($F_{8,62}$ = 1.48 P=0.18), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Caenimonas* defined by RDP classification at deepest assignment.

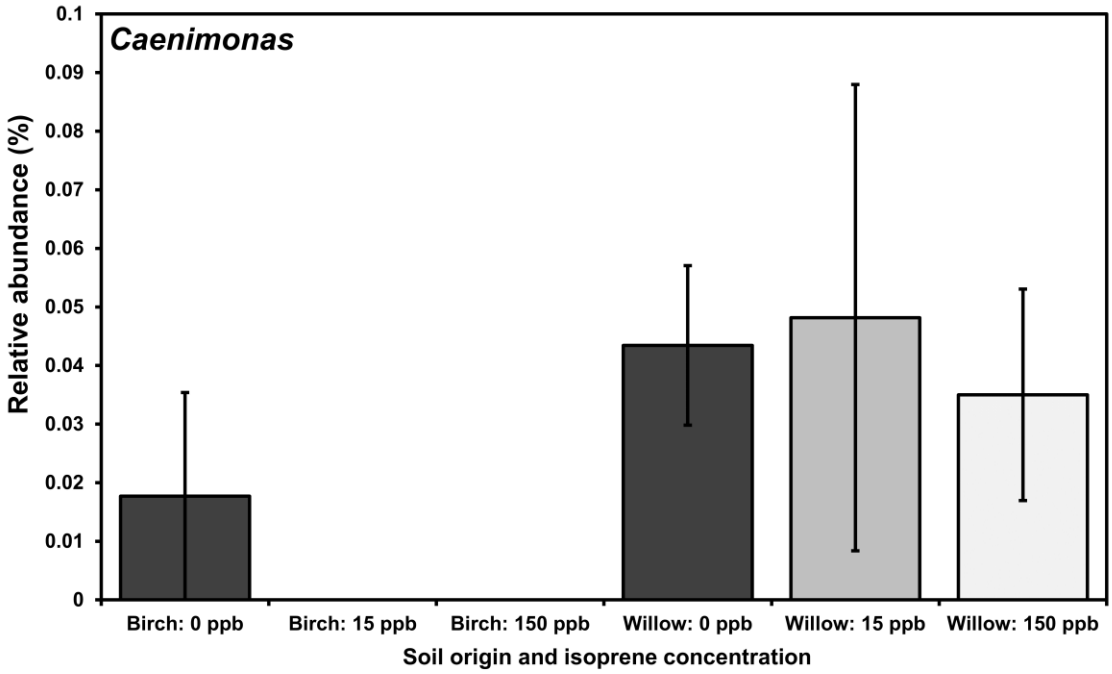


Figure 2.SI.64 *Caenimonas* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Caenimonas* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 0.66 P=0.73) and was not affected by the tree type ($F_{8,15}$ = 1.89 P=0.14).

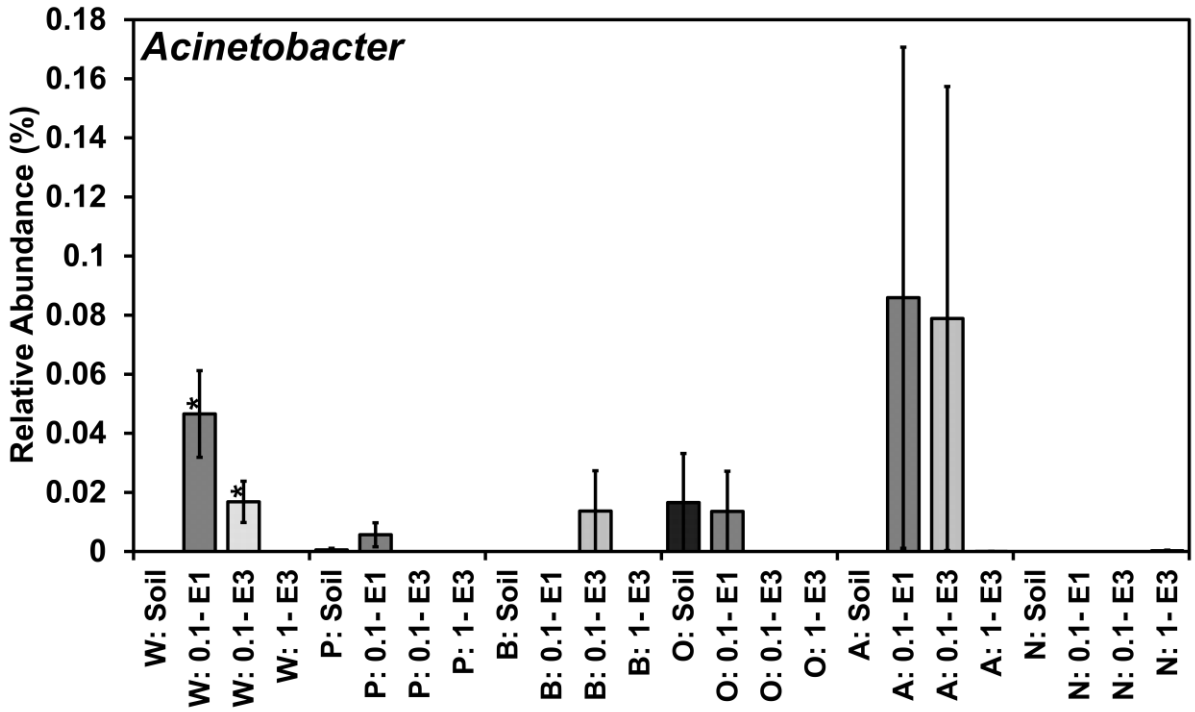


Figure 2.SI.65 *Acinetobacter* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was affected by sample location ($F_{8,62}$ = 2.1 P=0.05), not affected by concentration ($F_{8,62}$ = 1.49 P=0.18), and was not affected by enrichment level ($F_{8,62}$ = 0.56 P=0.81), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Acinetobacter* defined by RDP classification at deepest assignment.

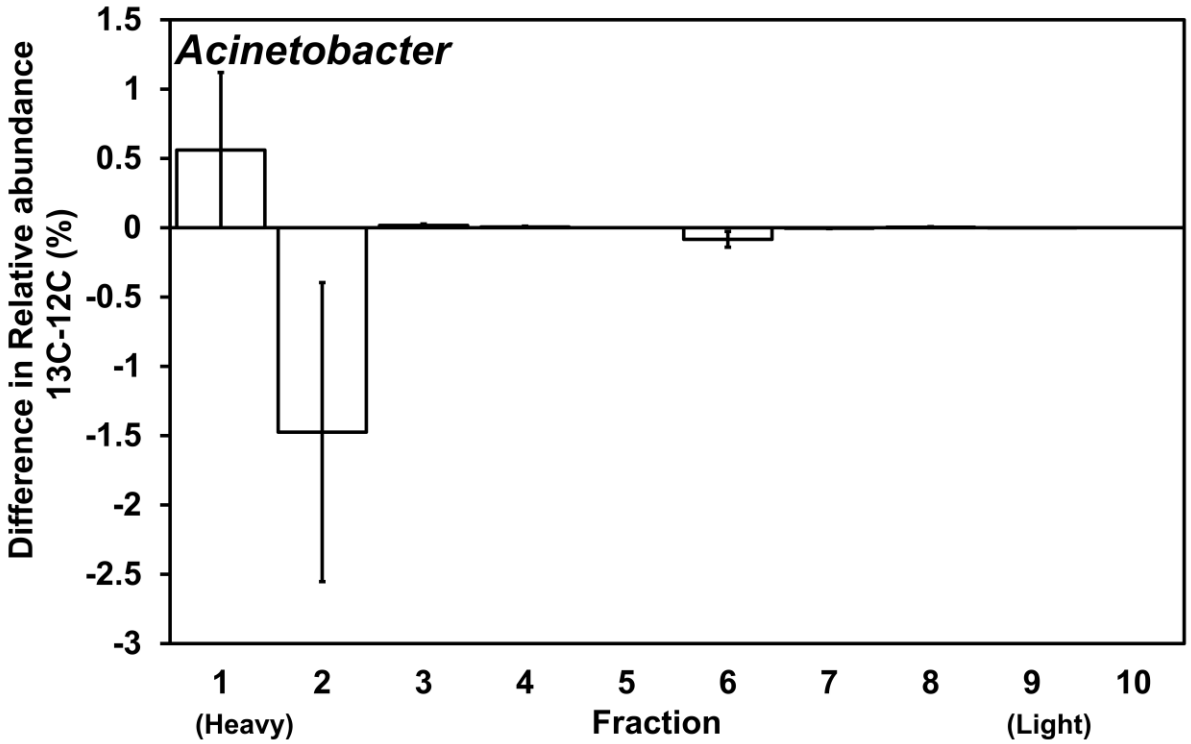


Figure 2.SI.66 *Acinetobacter* relative abundances in ^{13}C fractions net of *Acinetobacter* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Acinetobacter* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was significantly affected by the density ($F_{3,65}$ = 4.31 P=0.01), and was not affected by the carbon type ($F_{3,65}$ = 1.11 P=0.36), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

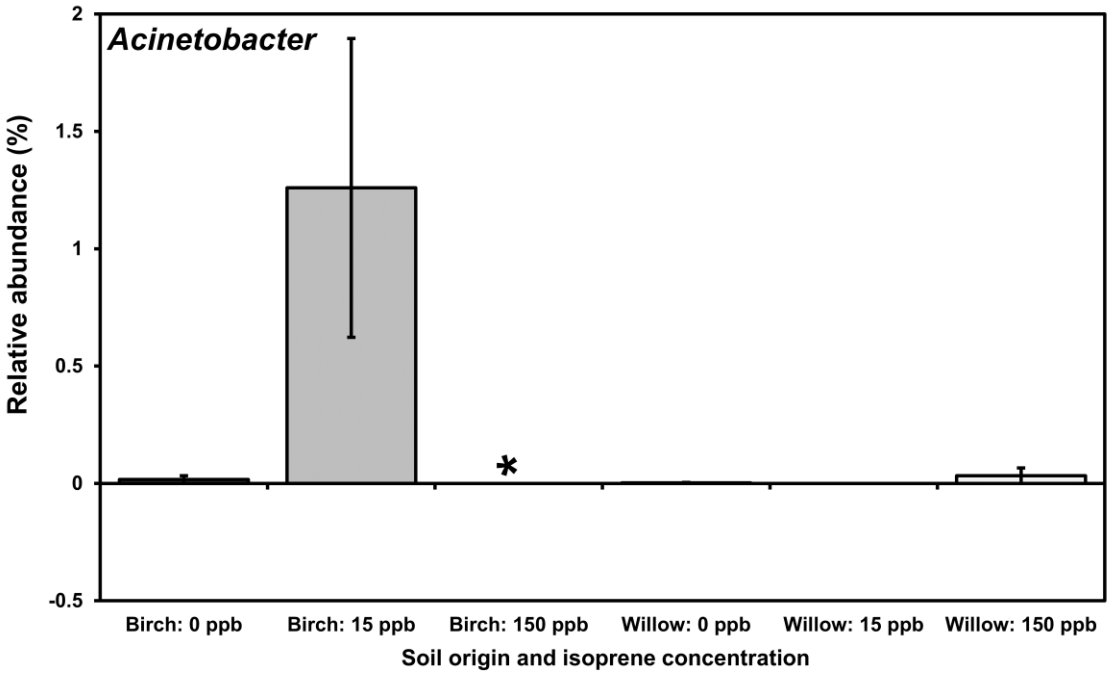


Figure 2.SI.67 *Acinetobacter* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Acinetobacter* defined by RDP classification at deepest assignment.

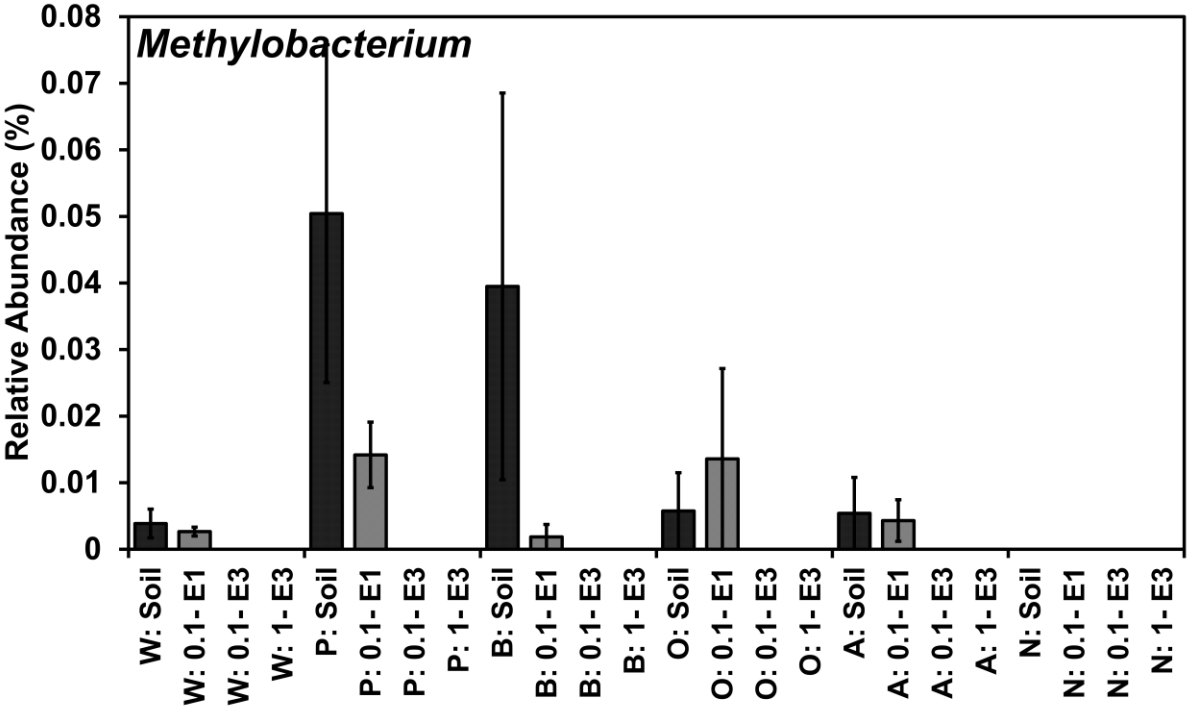


Figure 2.SI.68 *Methylobacterium* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated

isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 0.91 P=0.52), not affected by concentration ($F_{8,62}$ = 0.32 P=0.96), and was not affected by enrichment level ($F_{8,62}$ = 1.9 P=0.08), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Methylobacterium* defined by RDP classification at deepest assignment.

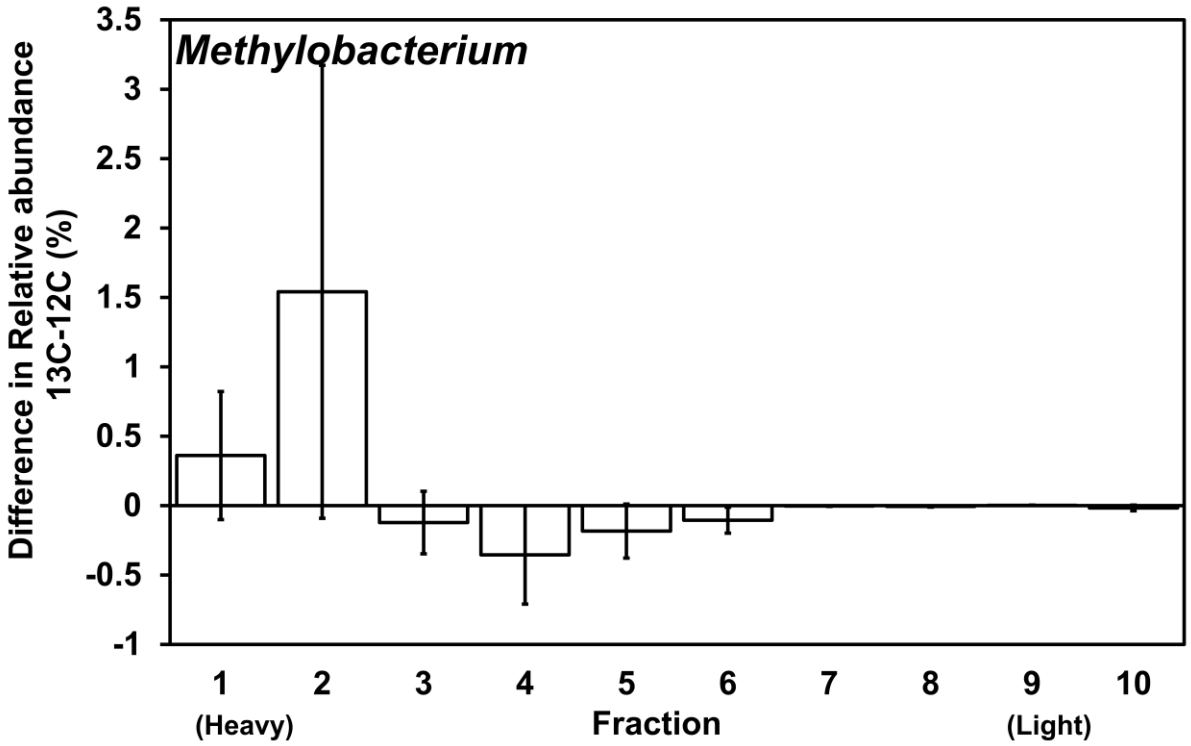


Figure 2.SI.69 *Methylobacterium* relative abundances in ^{13}C fractions net of *Methylobacterium* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Methylobacterium* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was significantly affected by the density ($F_{3,65}$ = 2.72 P=0.05), and was not affected by the carbon type ($F_{3,65}$ = 0.77 P=0.52), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

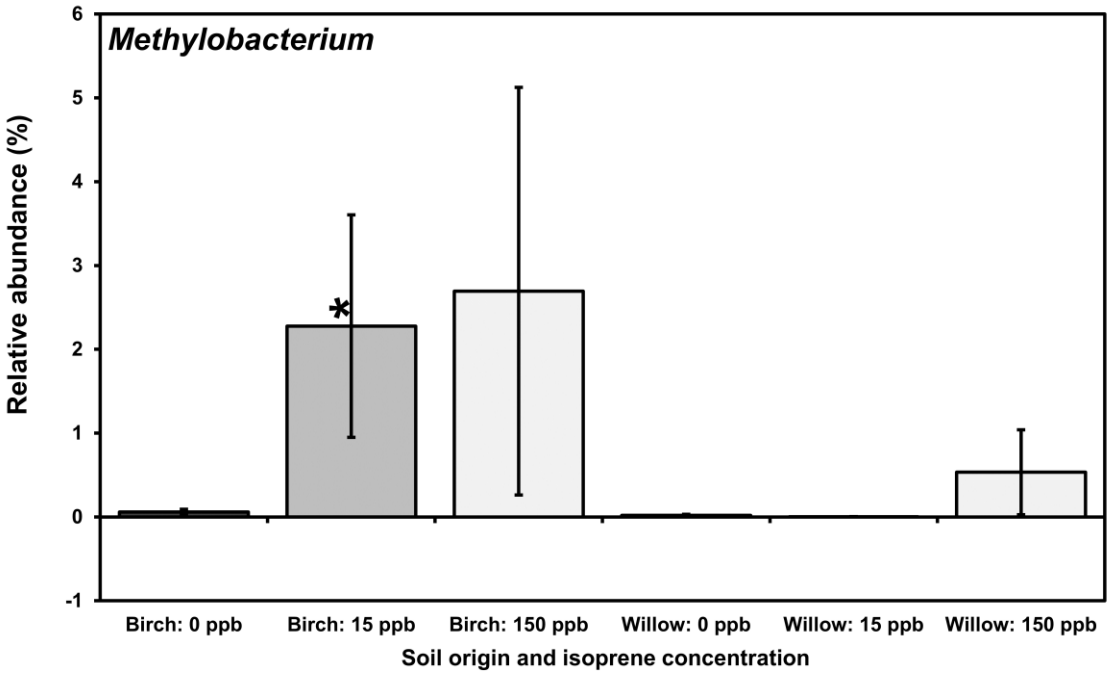


Figure 2.SI.70 *Methylobacterium* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Methylobacterium* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 1.07 P=0.43) and was not affected by the tree type ($F_{8,15}$ = 0.27 P=0.97).

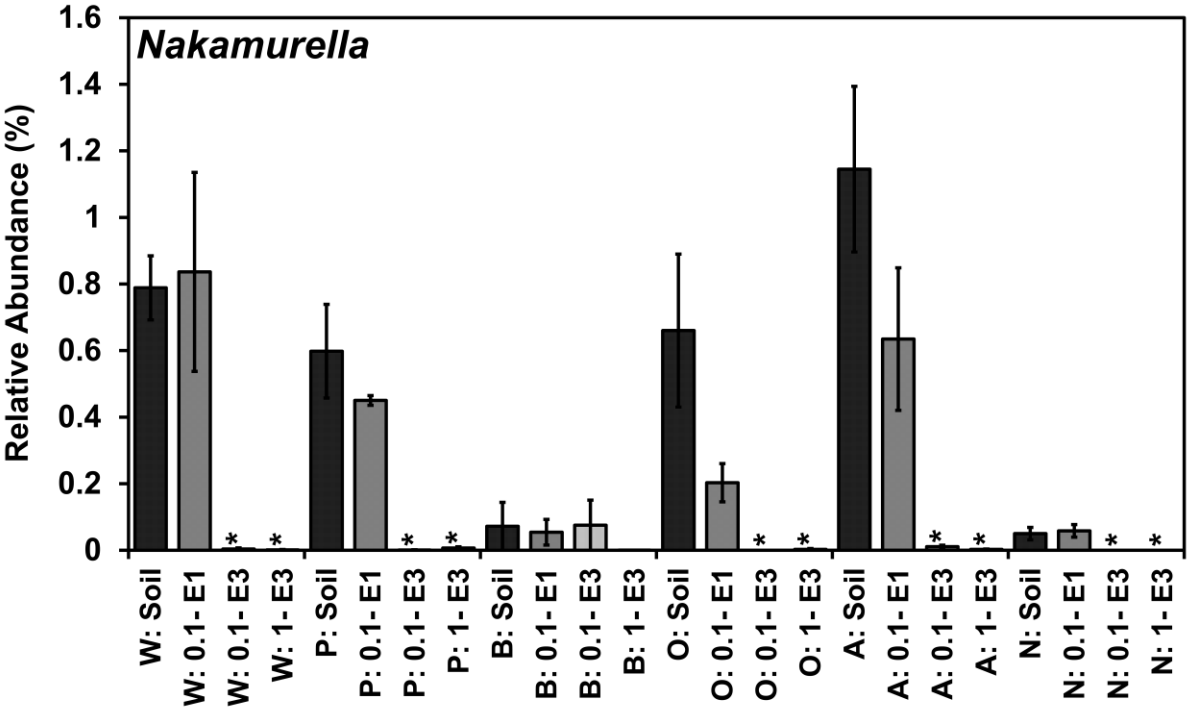


Figure 2.SI.71 *Nakamurella* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was not affected by sample location ($F_{8,62}$ = 1.55 P=0.16), not affected by concentration ($F_{8,62}$ = 0.6 P=0.78), and was affected by enrichment level ($F_{8,62}$ = 2.08 P=0.05), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only), n = 3, Error bars = SE, *Nakamurella* defined by RDP classification at deepest assignment.

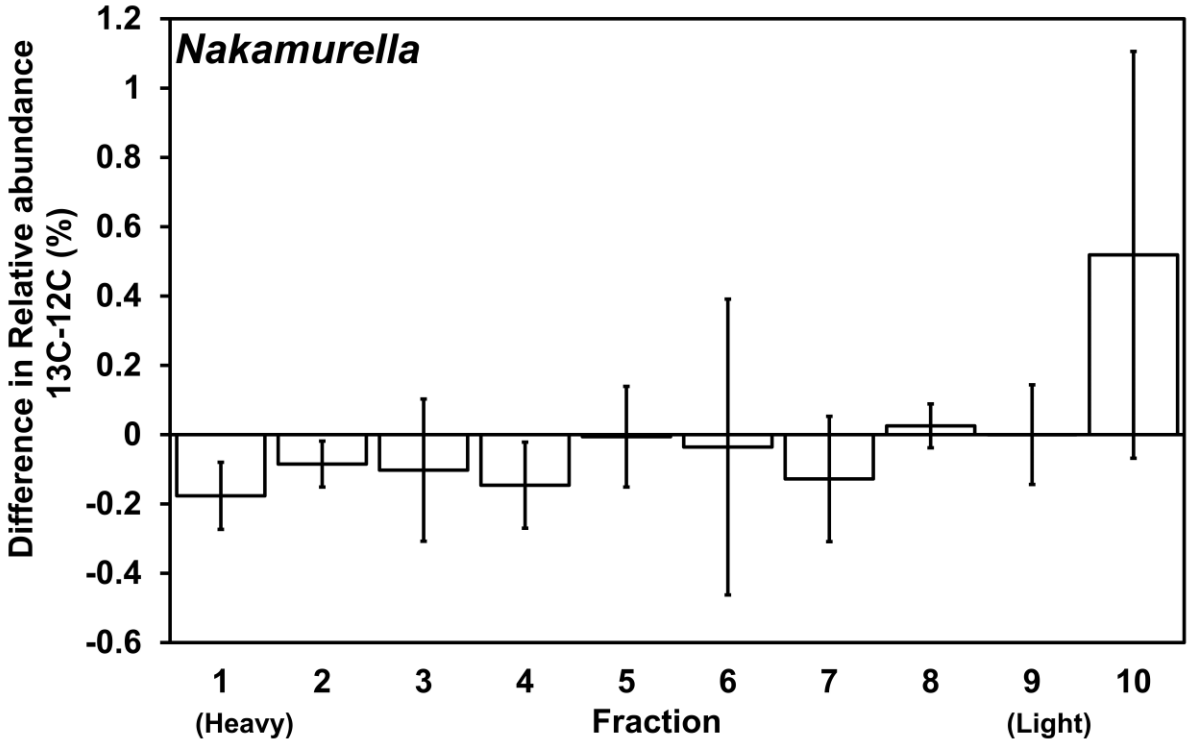


Figure 2.SI.72 *Nakamurella* relative abundances in ^{13}C fractions net of *Nakamurella* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Nakamurella* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that the abundance was significantly affected by the density ($F_{3,65}$ = 3.8 P=0.01), and was not affected by the carbon type ($F_{3,65}$ = 0.3 P=0.83), * = significance at p<0.05 (in univariate unadjusted Kruskal-Wallis tests, for display purposes only).

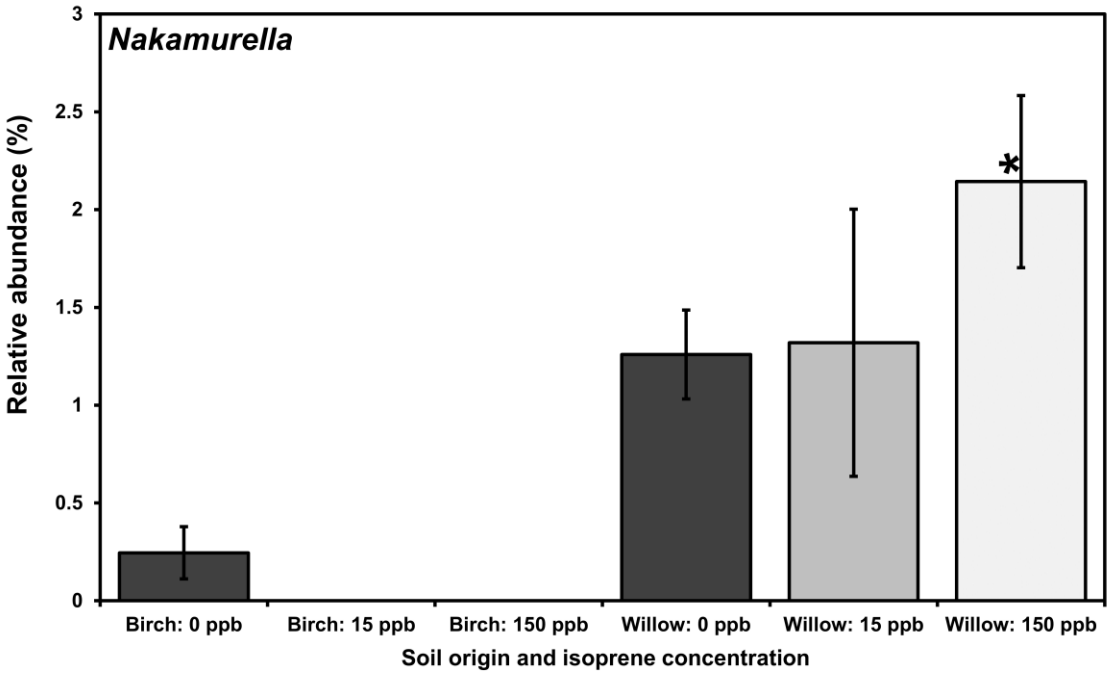


Figure 2.SI.73 *Nakamurella* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Nakamurella* defined by RDP classification at deepest assignment. A multi-variable model fitted to the data set using manyglm (with a negative bionomial distribution assumption) within the R package mvabund shown that abundance was not affected by the isoprene concentration ($F_{8,15}$ = 0.63 P=0.74) and was affected by the tree type ($F_{8,15}$ = 2.72 P=0.05).

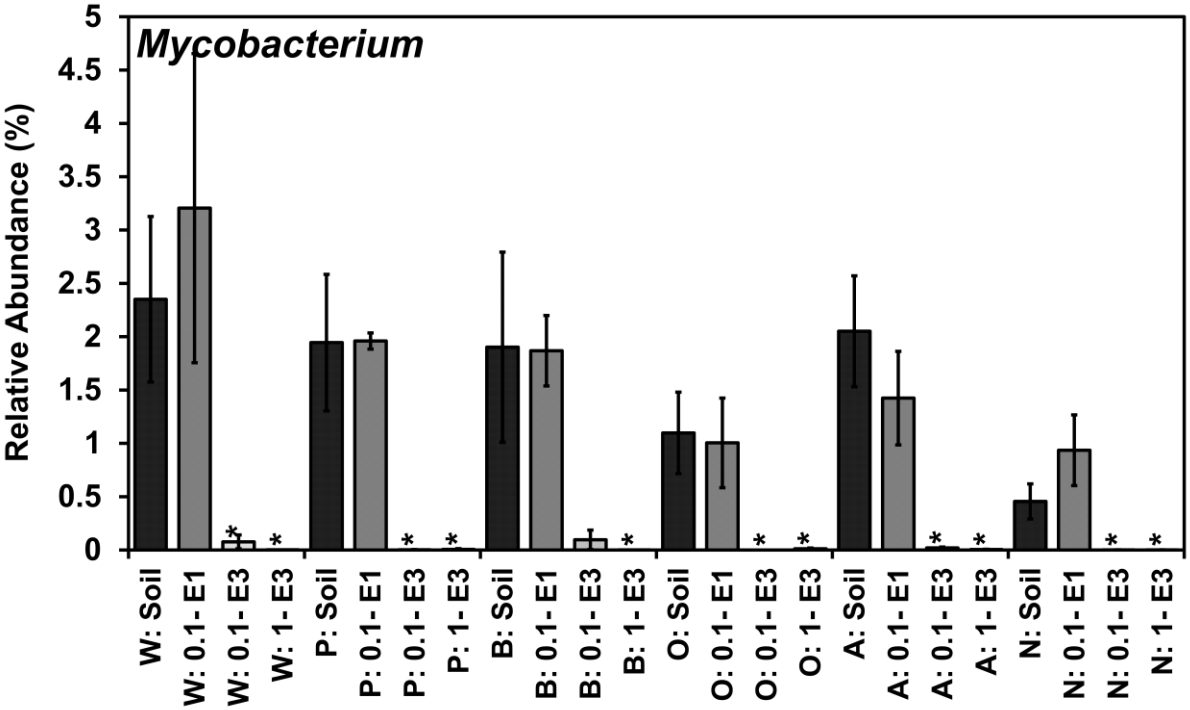


Figure 2.SI.74 *Mycobacterium* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Mycobacterium* defined by RDP classification at deepest assignment.

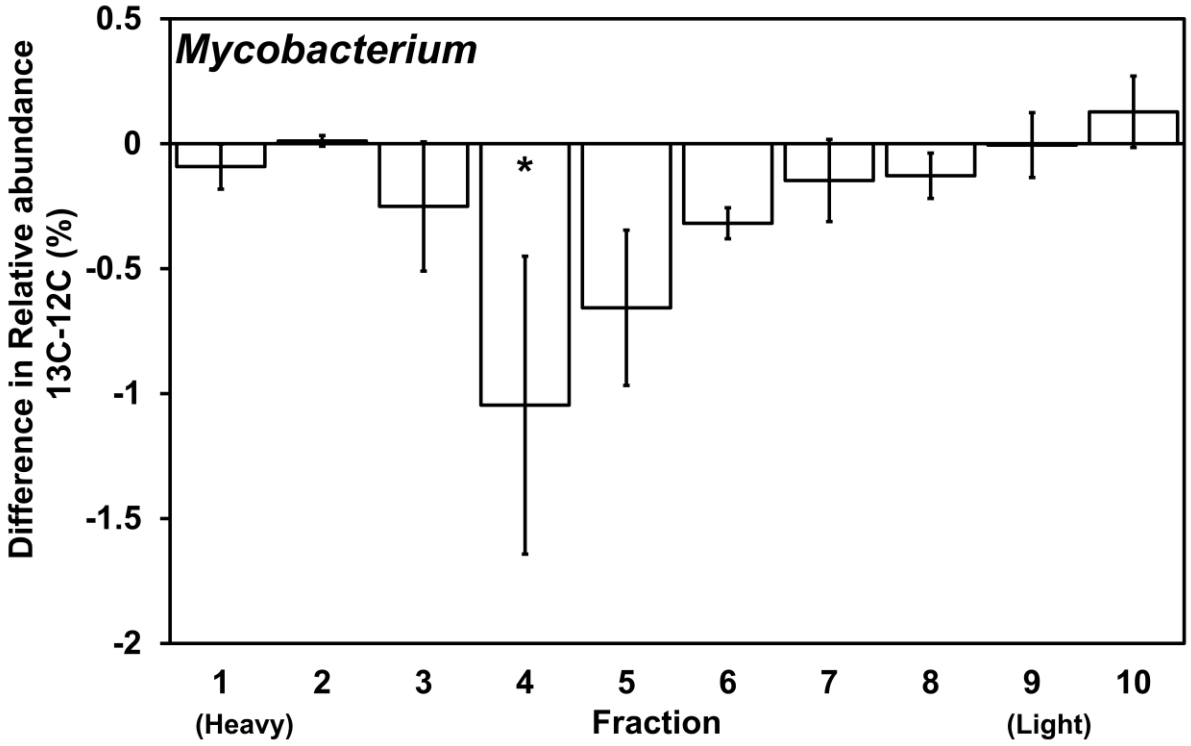


Figure 2.SI.75 *Mycobacterium* relative abundances in ^{13}C fractions net of *Mycobacterium* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Mycobacterium* defined by RDP classification at deepest assignment.

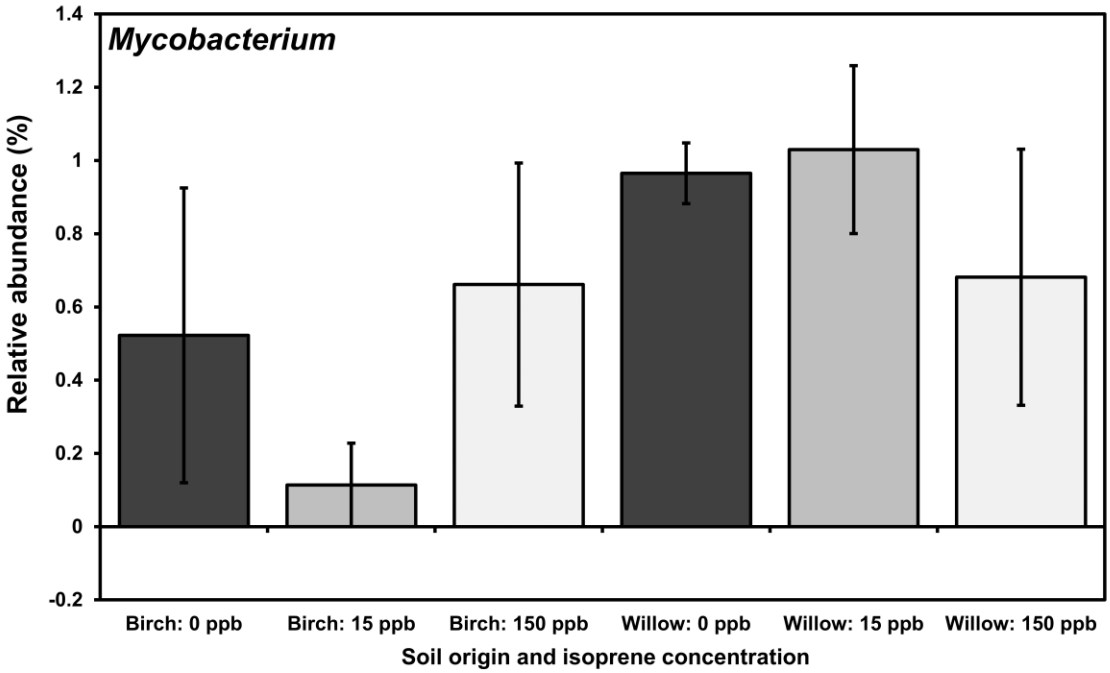


Figure 2.SI.76 *Mycobacterium* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Mycobacterium* defined by RDP classification at deepest assignment.

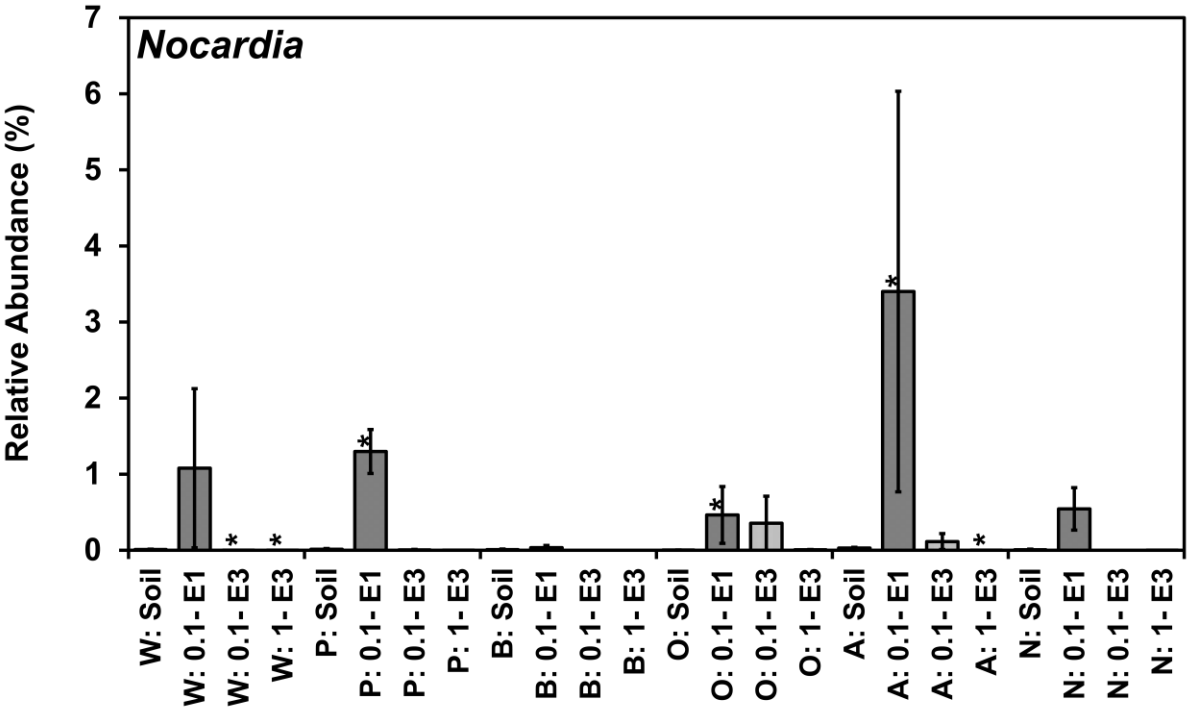


Figure 2.SI.77 *Nocardia* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second

enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Nocardia* defined by RDP classification at deepest assignment.

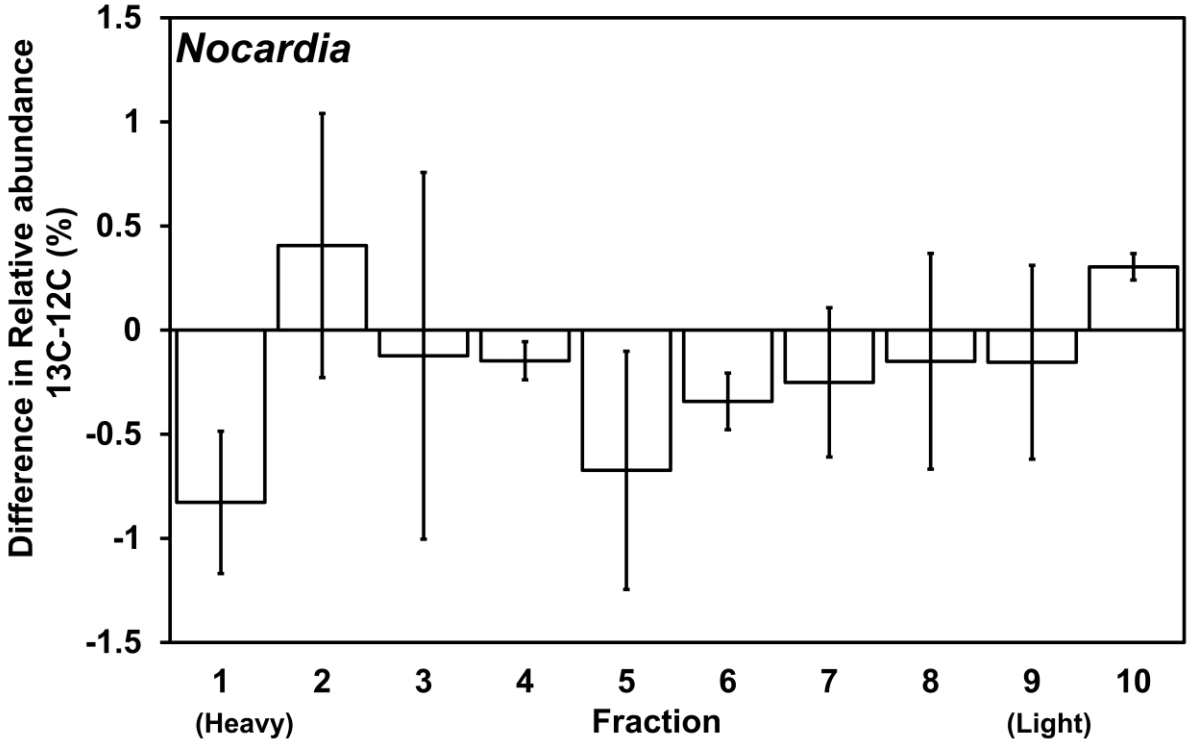


Figure 2.SI75 *Nocardia* relative abundances in ^{13}C fractions net of *Nocardia* relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Nocardia* defined by RDP classification at deepest assignment.

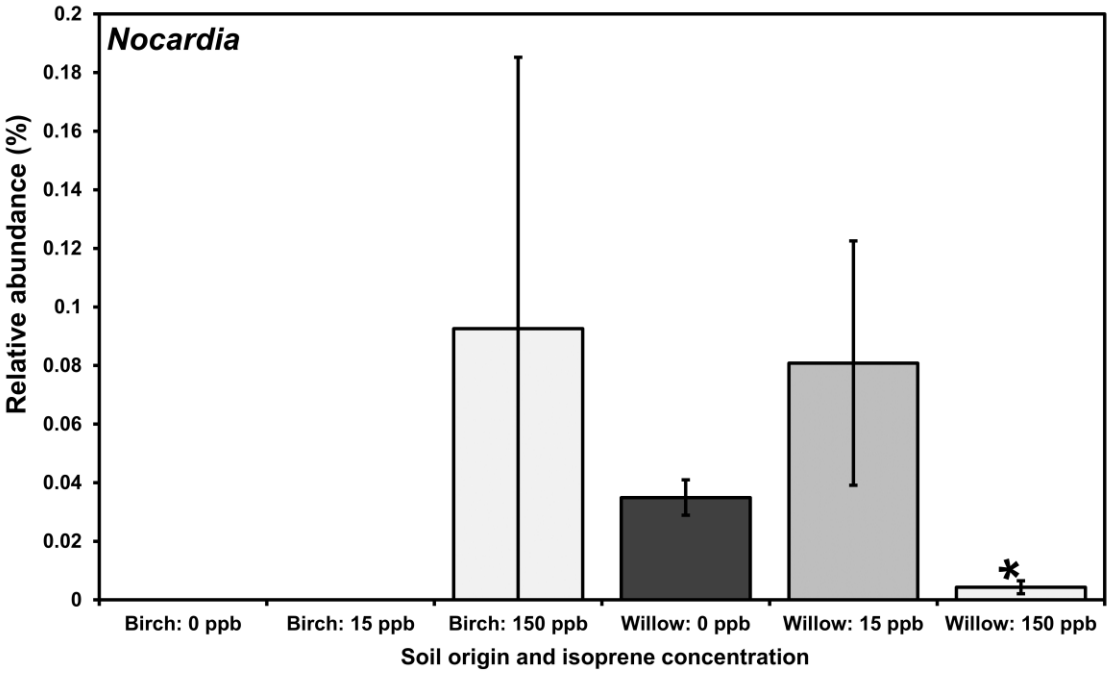


Figure 2.SI.79 *Nocardia* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Nocardia* defined by RDP classification at deepest assignment.

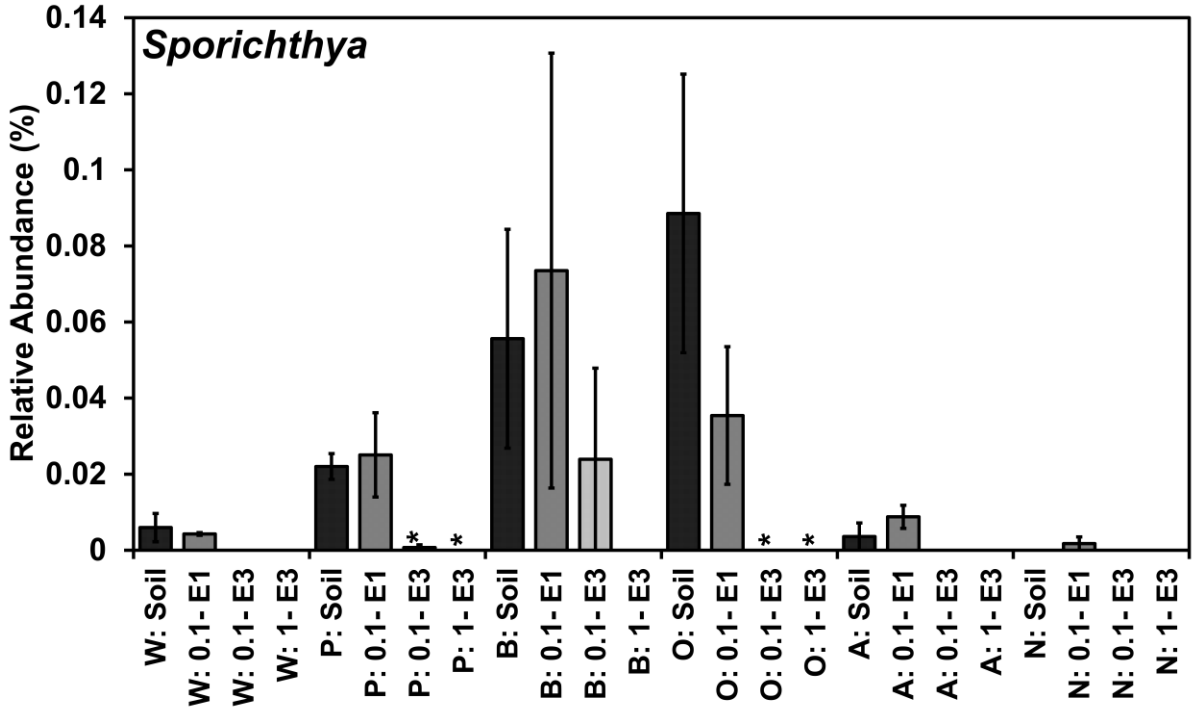


Figure 2.SI.80 *Sporichthya* relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at $p<0.05$ (Kruskal-Wallis), n = 3, Error bars = SE, *Sporichthya* defined by RDP classification at deepest assignment.

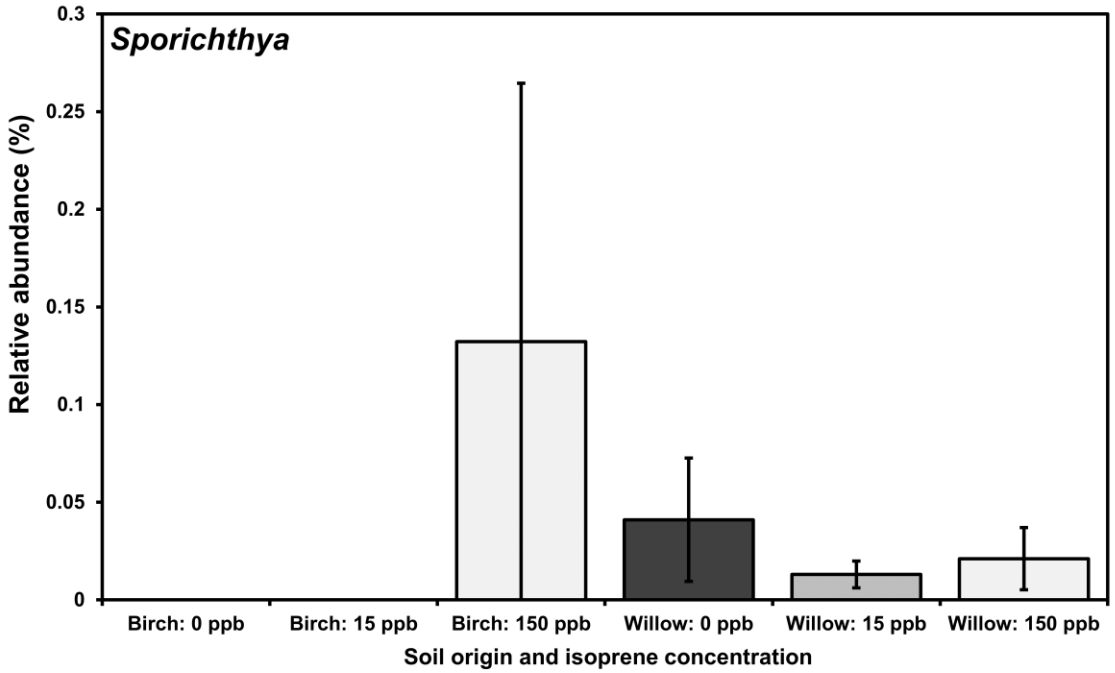


Figure 2.SI.81 *Sporichthya* relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Sporichthya* defined by RDP classification at deepest assignment.

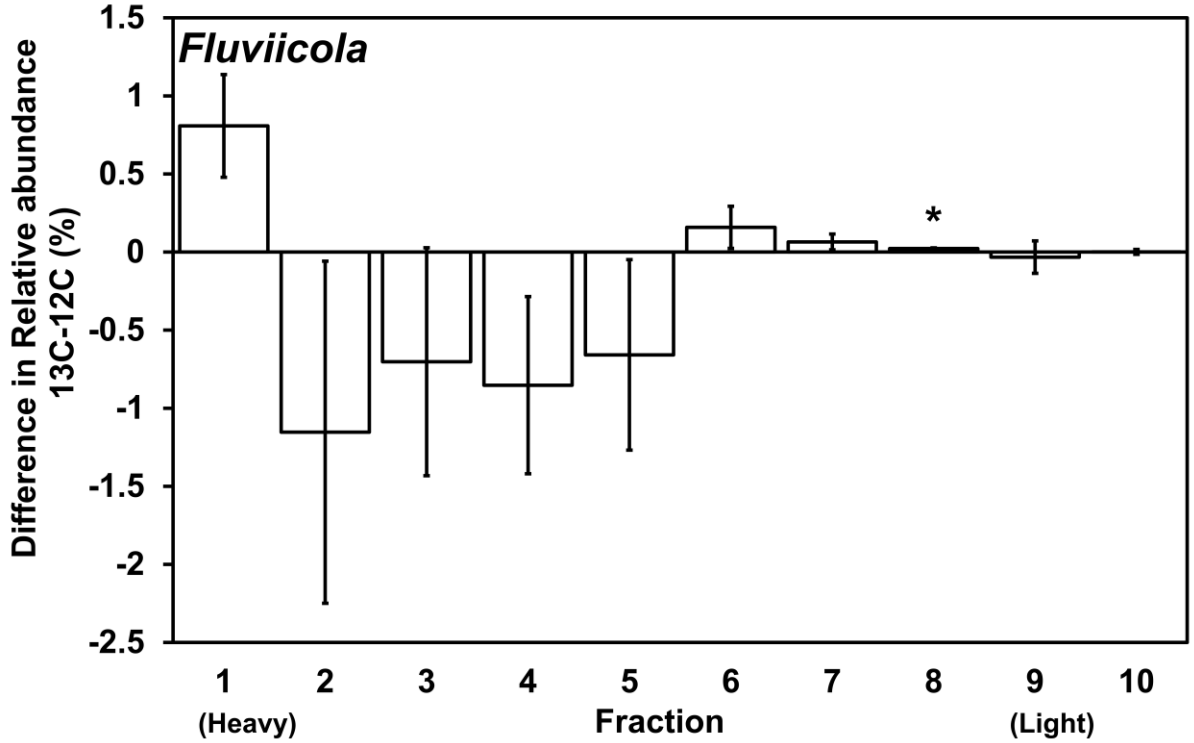


Figure 2.SI79 *Fluviicola* OTU s1 relative abundances in ^{13}C fractions net of *Fluviicola* OUT s1 relative abundances in corresponding ^{12}C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, ^{13}C and ^{12}C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Fluviicola* defined by RDP classification at deepest assignment.

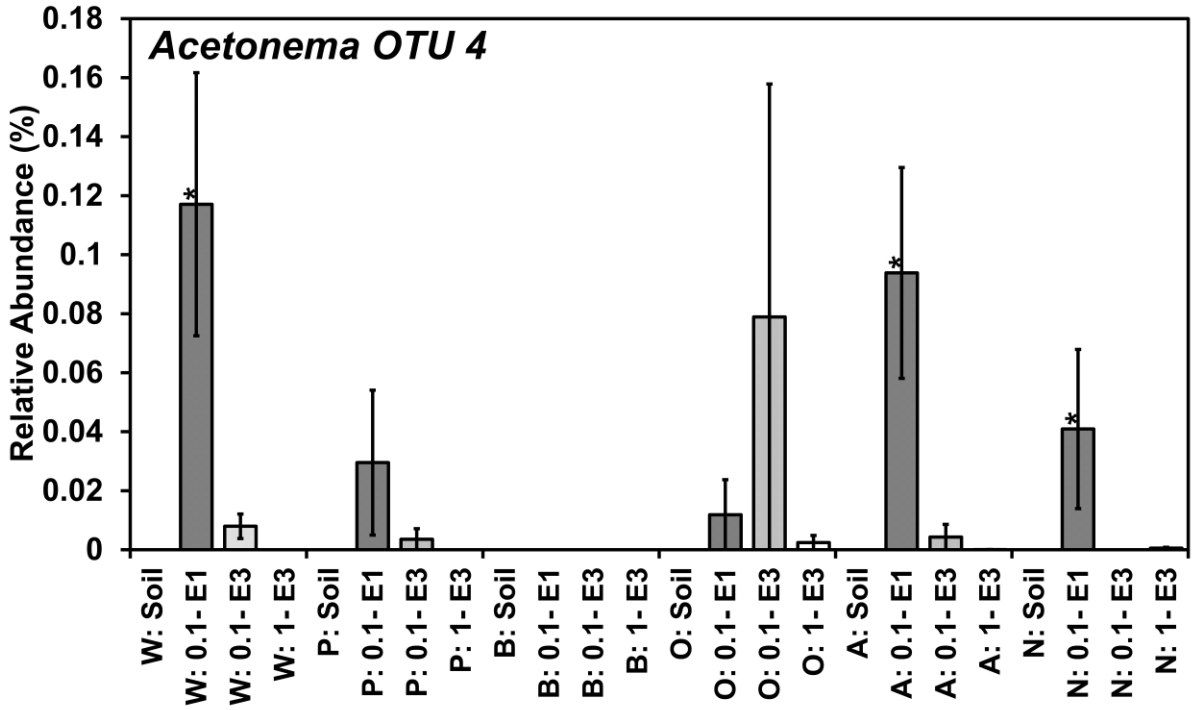


Figure 2.SI.82 *Acetonema* OTU 4 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Acetonema* OTU 4 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

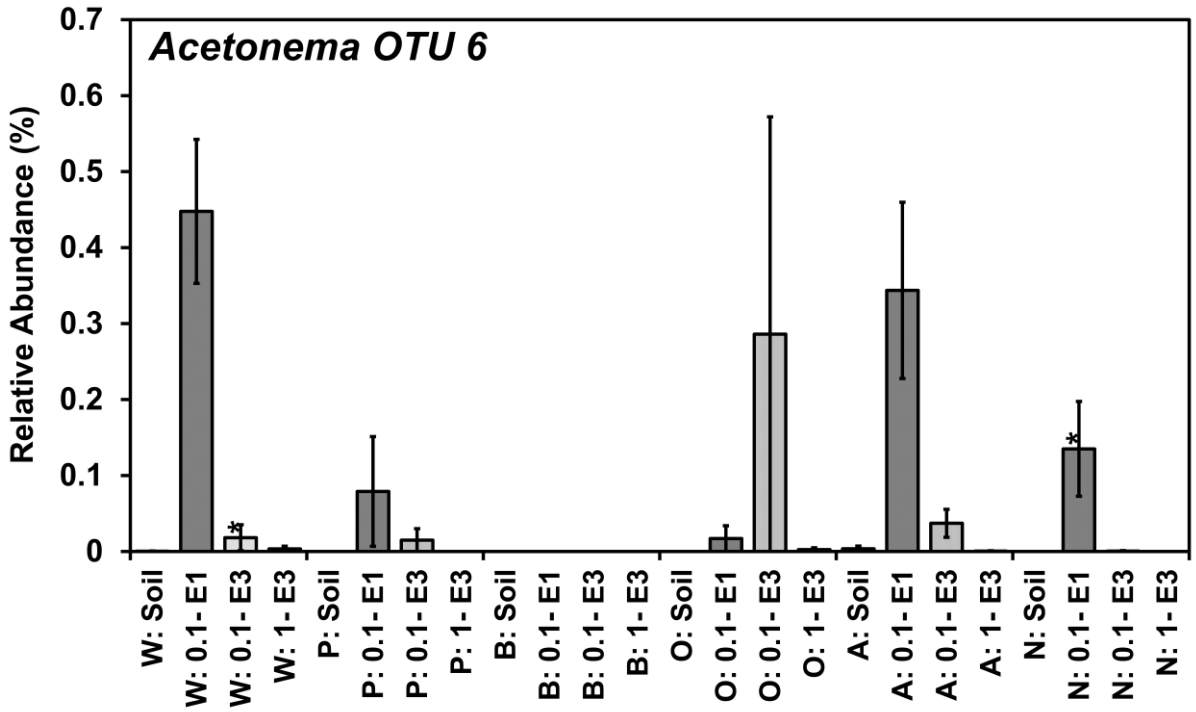


Figure 2.SI.83 *Acetonema* OTU 6 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Acetonema* OTU 6 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

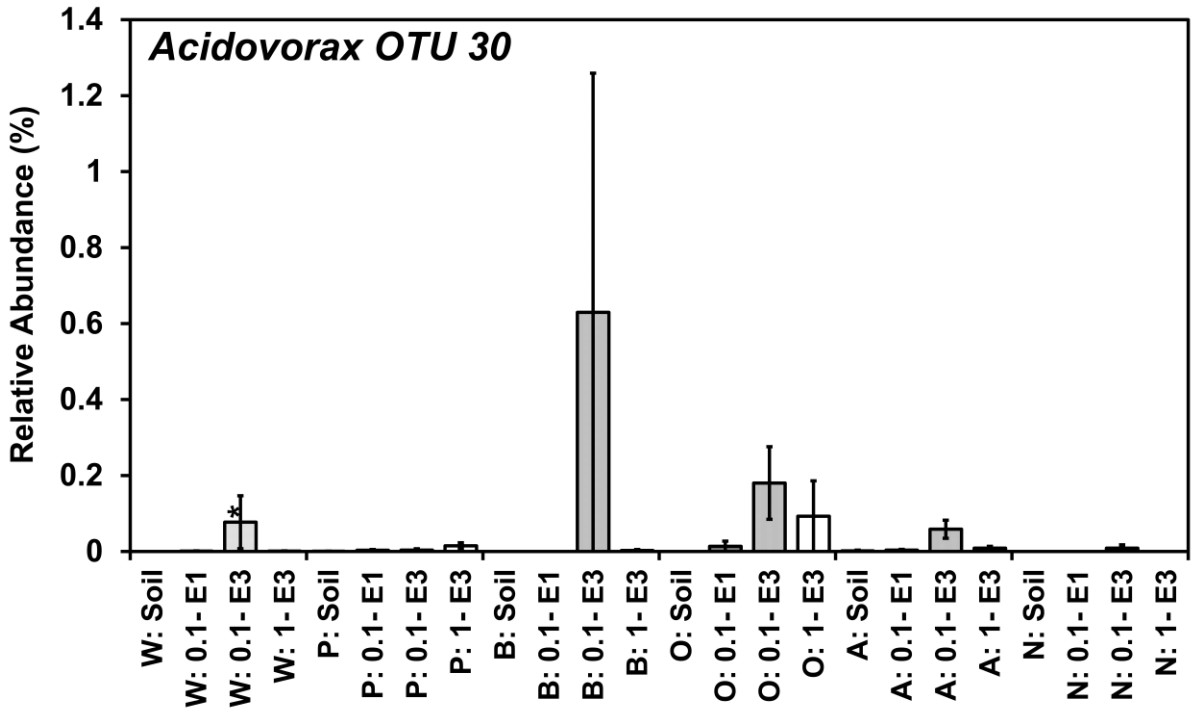


Figure 2.SI.84 *Acidovorax* OTU 30 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Acidovorax* OTU 30 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

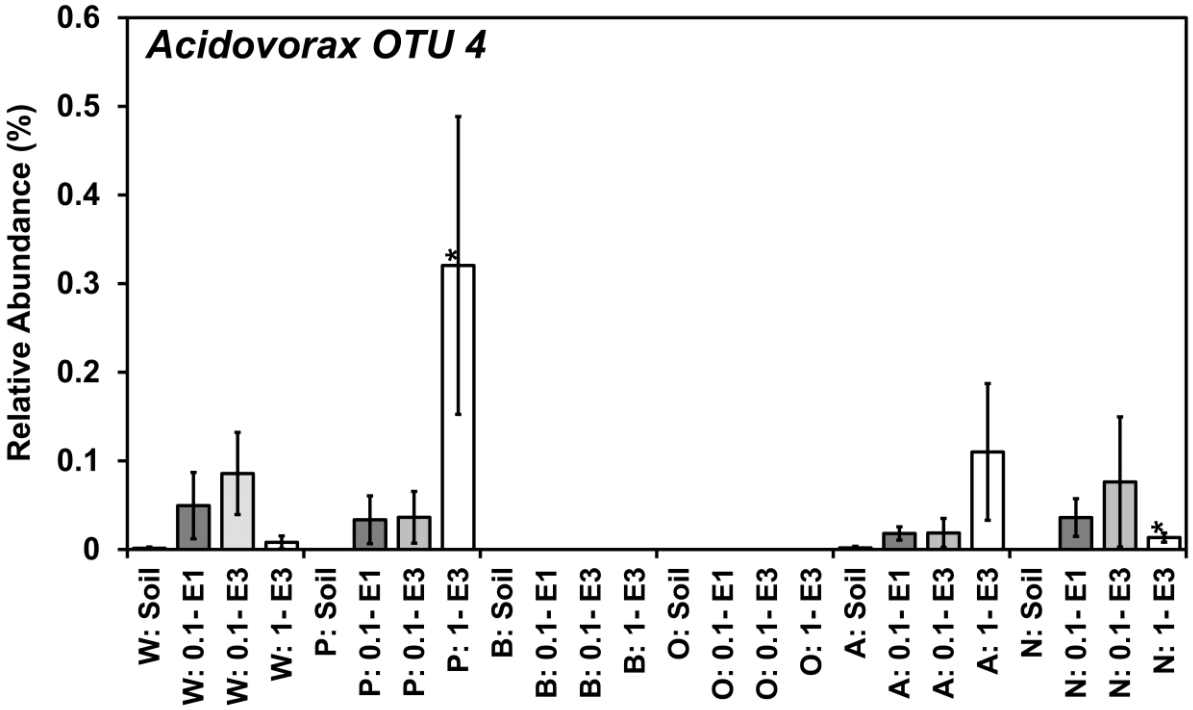


Figure 2.SI.85 *Acidovorax* OTU 4 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Acidovorax* OTU 4 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

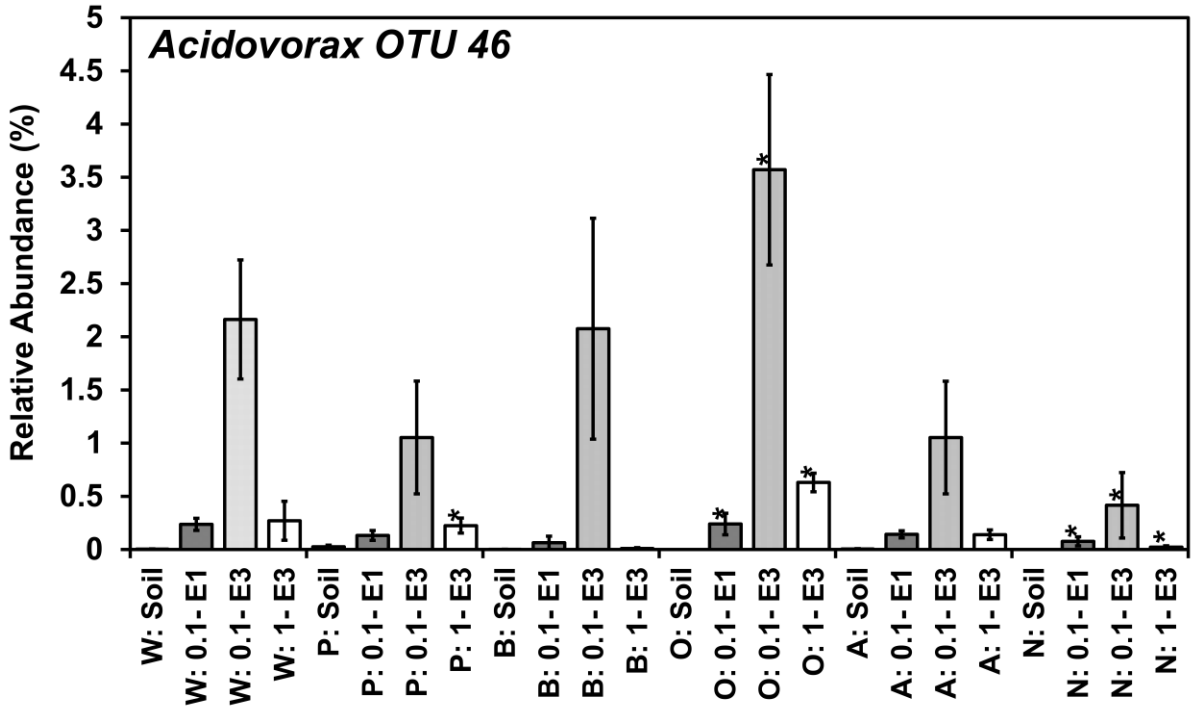


Figure 2.SI.86 *Acidovorax* OTU 46 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Acidovorax* OTU 46 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

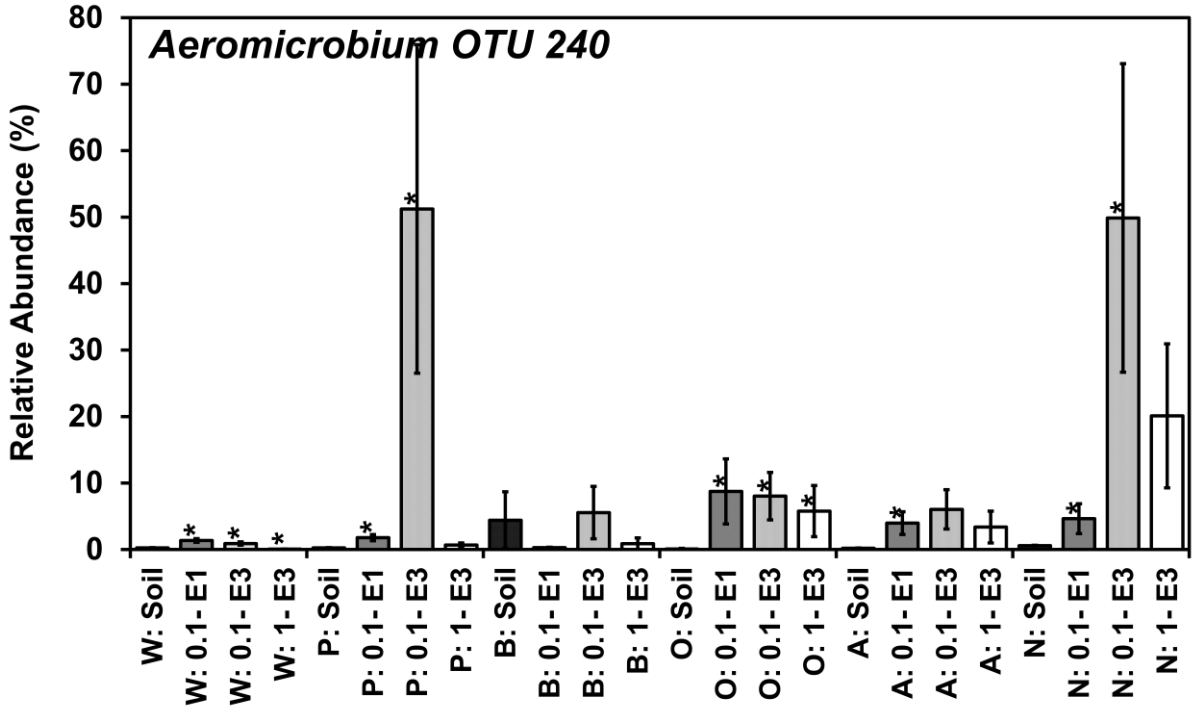


Figure 2.SI.87 *Aeromicrobium* OTU 240 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Aeromicrobium* OTU 240 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

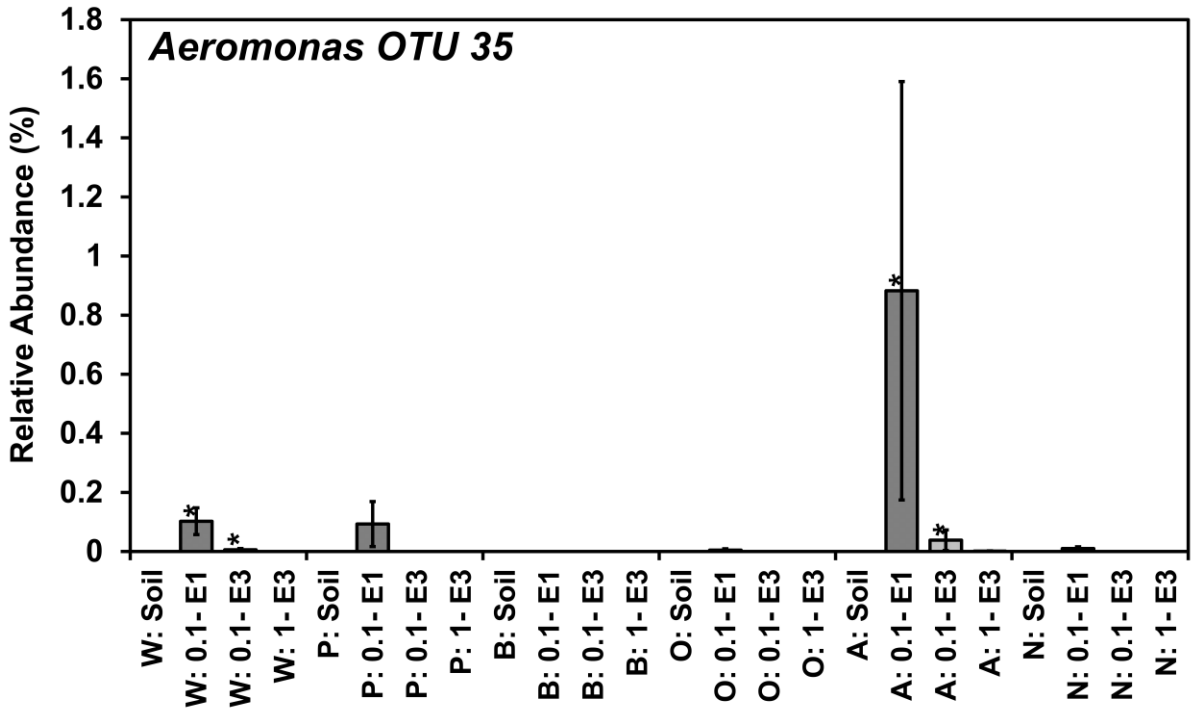


Figure 2.SI.88 *Aeromonas* OTU 35 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Aeromonas* OTU 35 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

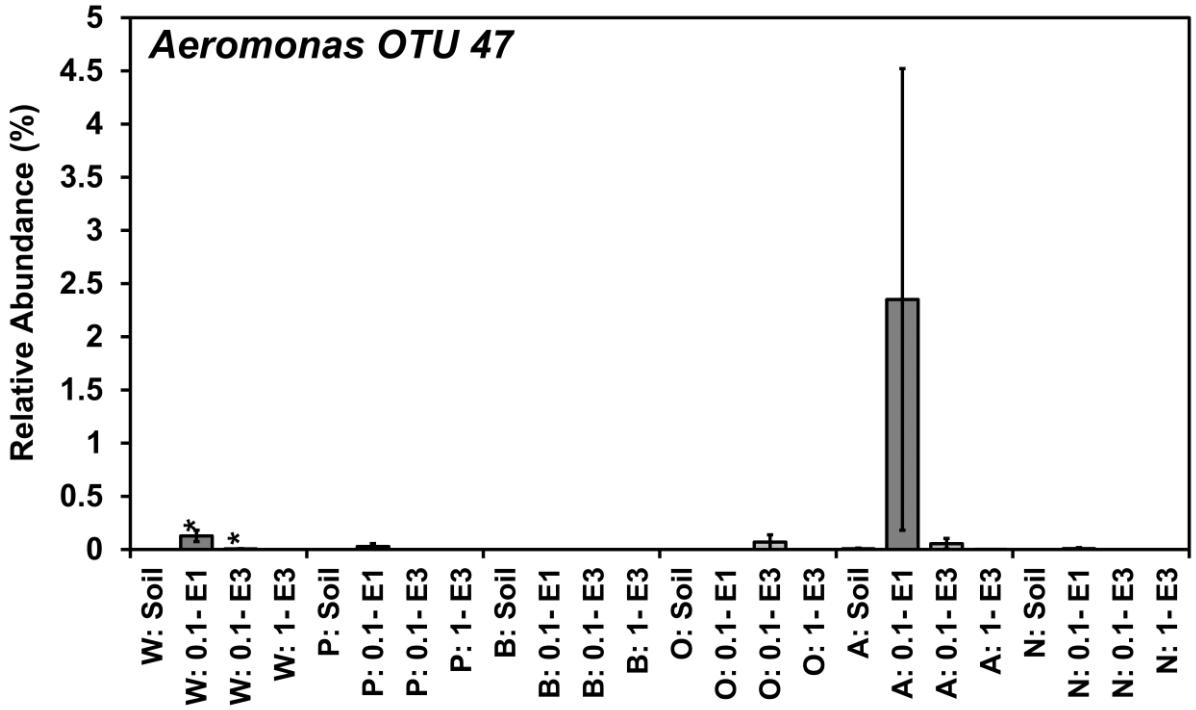


Figure 2.SI.89 *Aeromonas* OTU 47 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Aeromonas* OTU 47 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

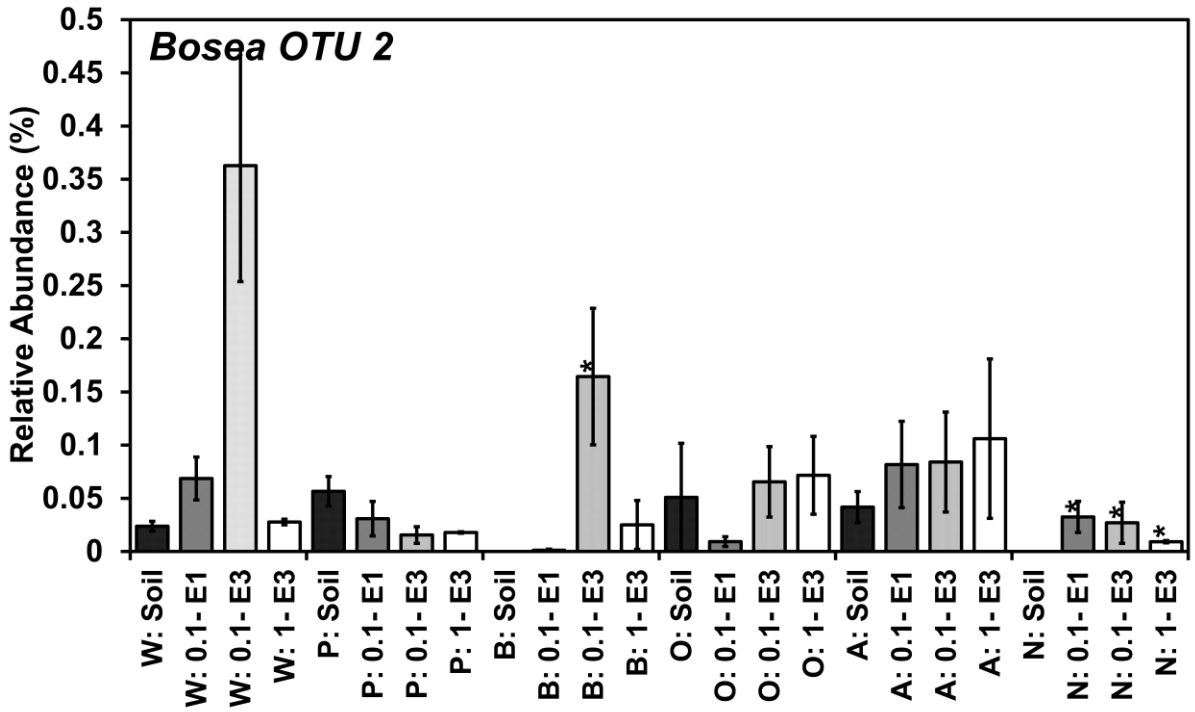


Figure 2.SI.90 *Bosea* OTU 2 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Bosea* OTU 2 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

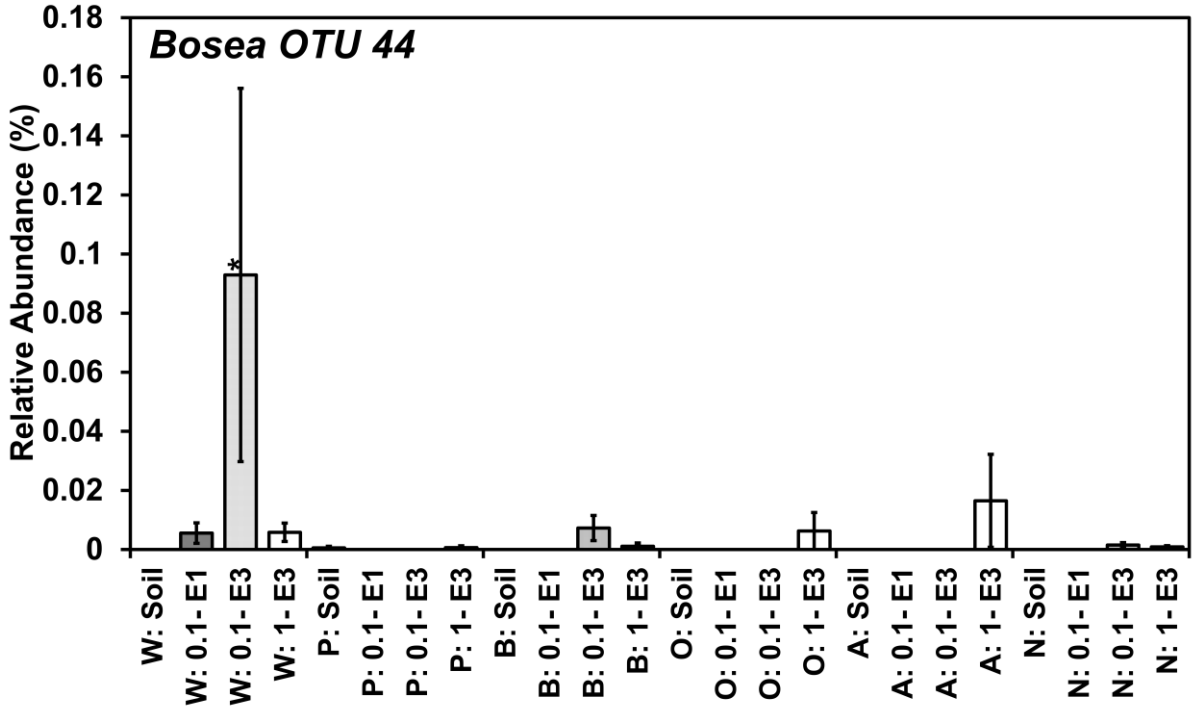


Figure 2.SI.91 *Bosea* OTU 44 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Bosea* OTU 44 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

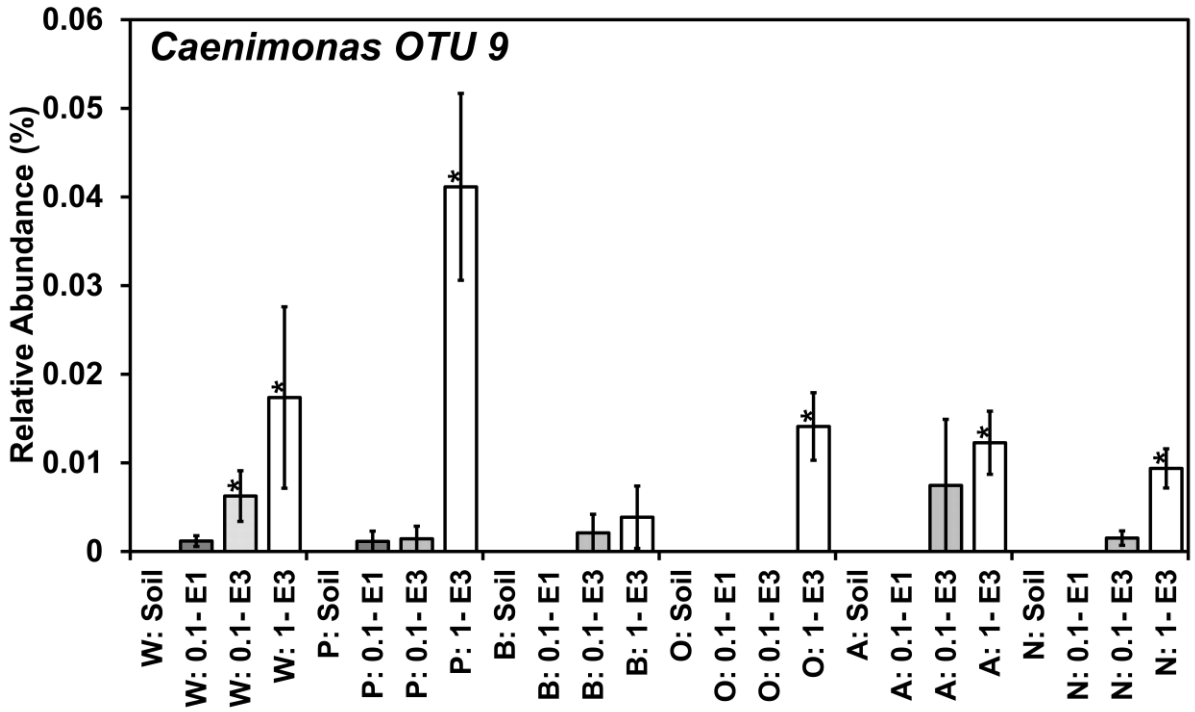


Figure 2.SI.92 *Caenimonas* OTU 9 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Caenimonas* OTU 9 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

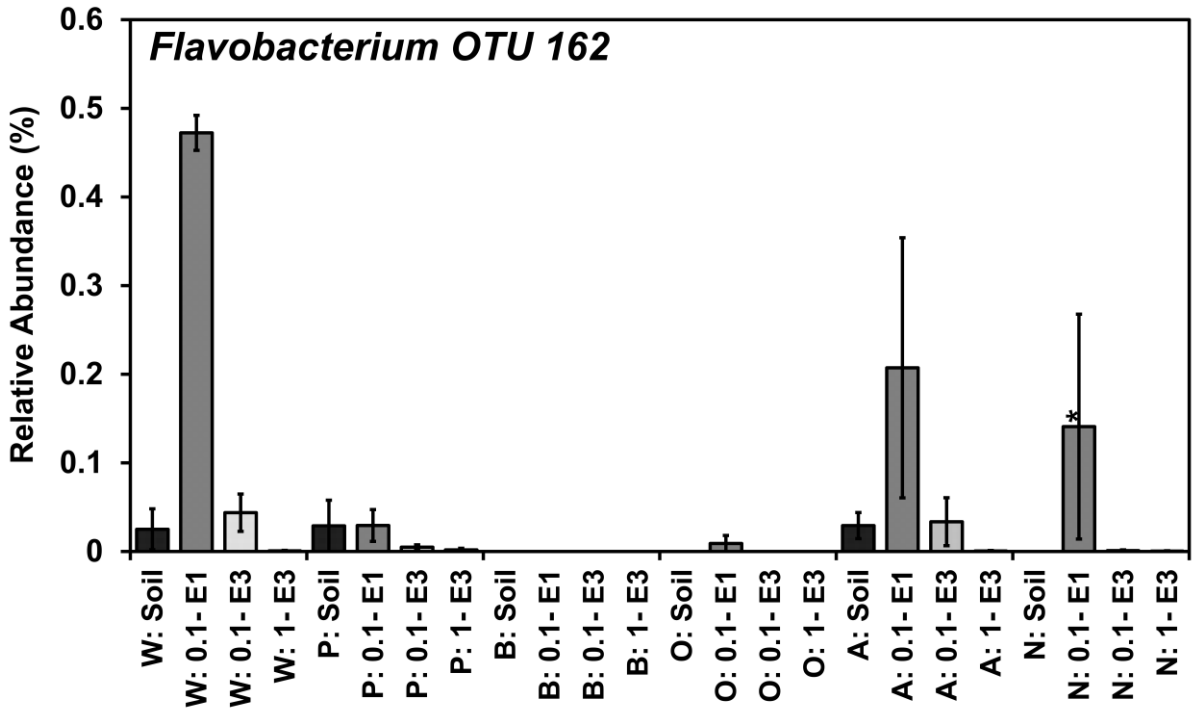


Figure 2.SI.93 *Flavobacterium* OTU 162 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Flavobacterium* OTU 162 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

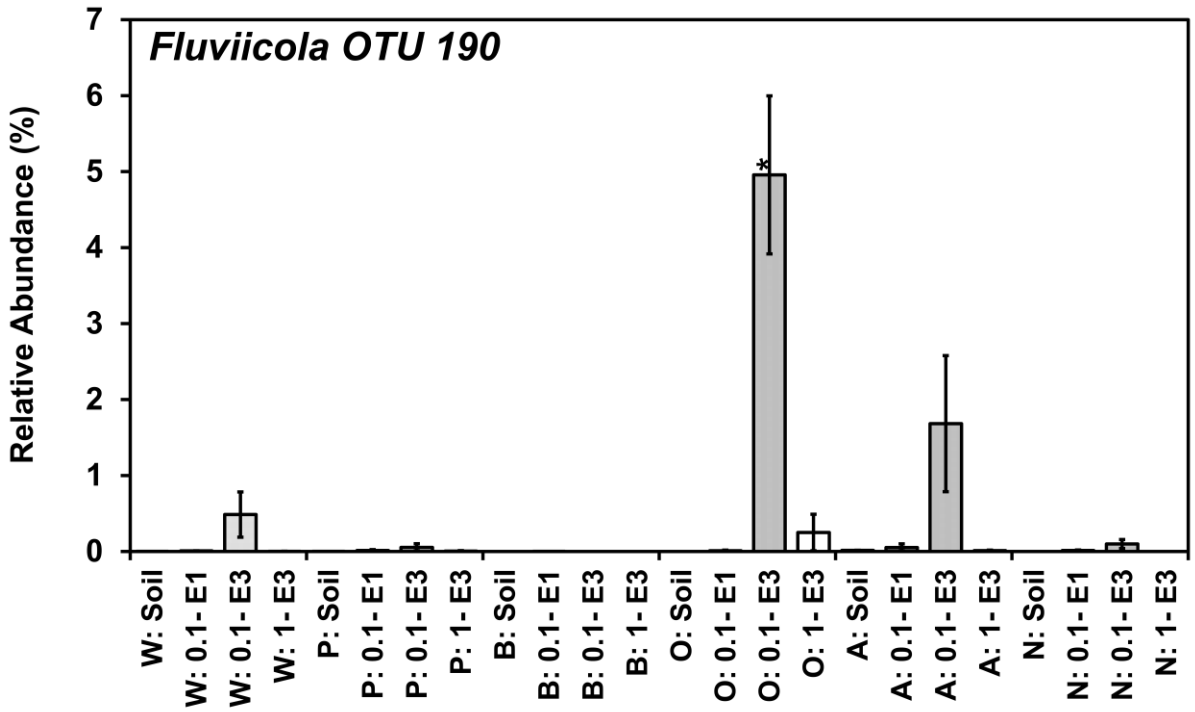


Figure 2.SI.94 *Fluviicola* OTU 190 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Fluviicola* OTU 190 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

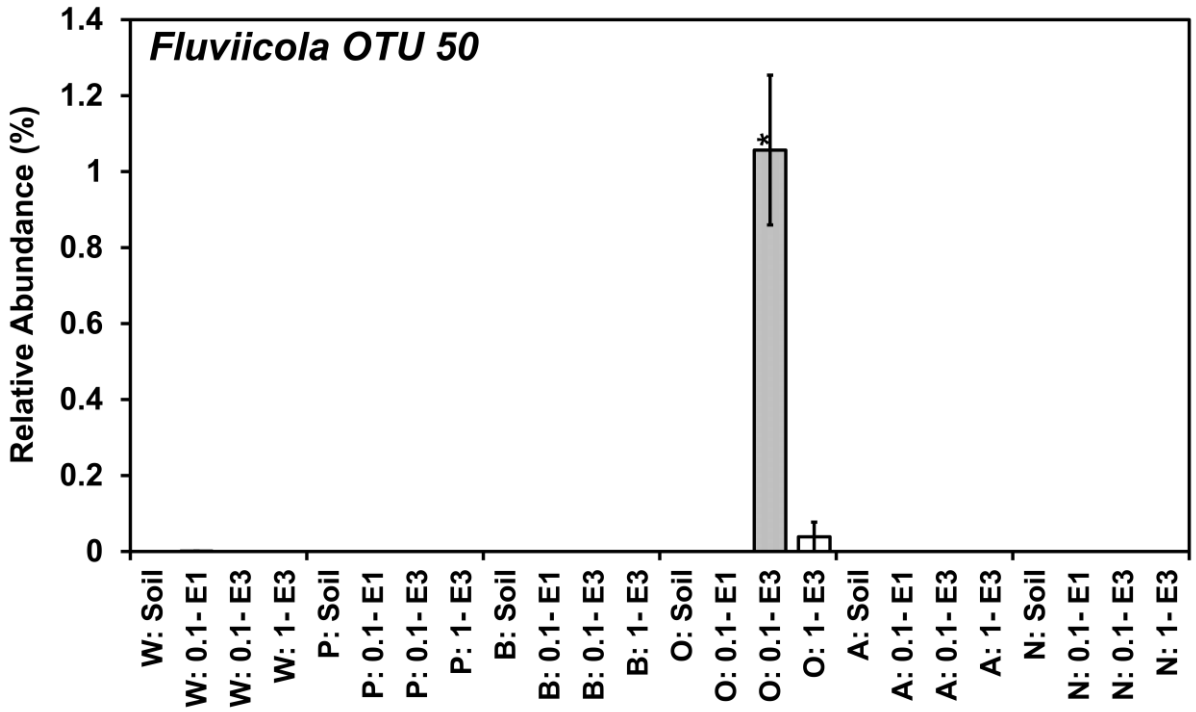


Figure 2.SI.95 *Fluviicola* OTU 50 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Fluviicola* OTU 50 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

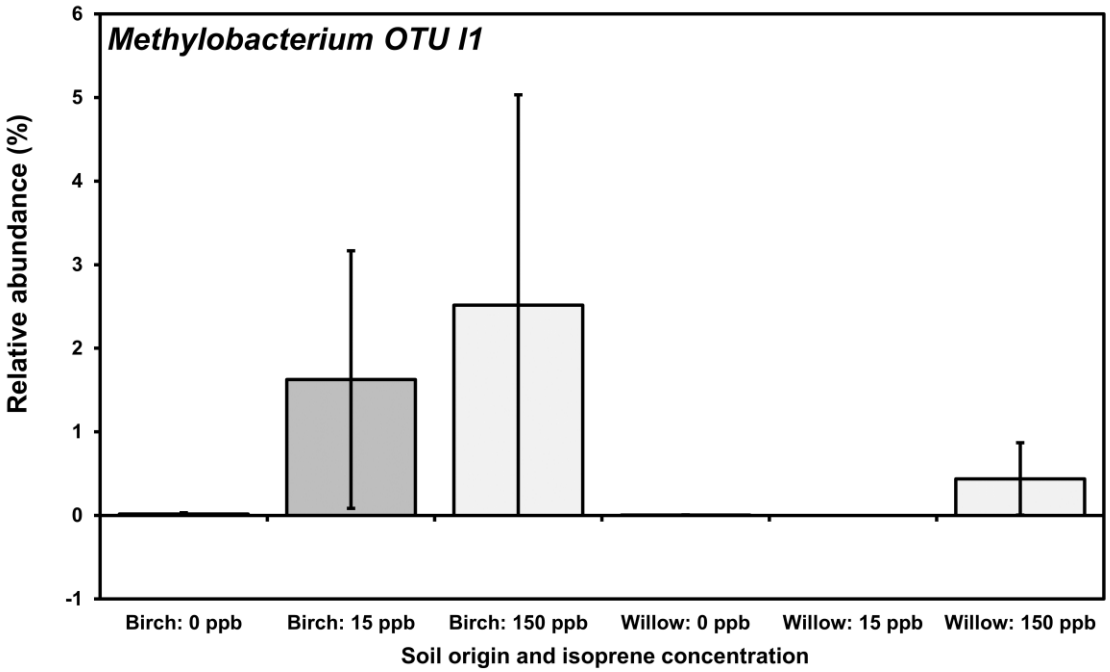


Figure 2.SI.96 *Methylobacterium* OTU l1 relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Methylobacterium* OTU l1 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

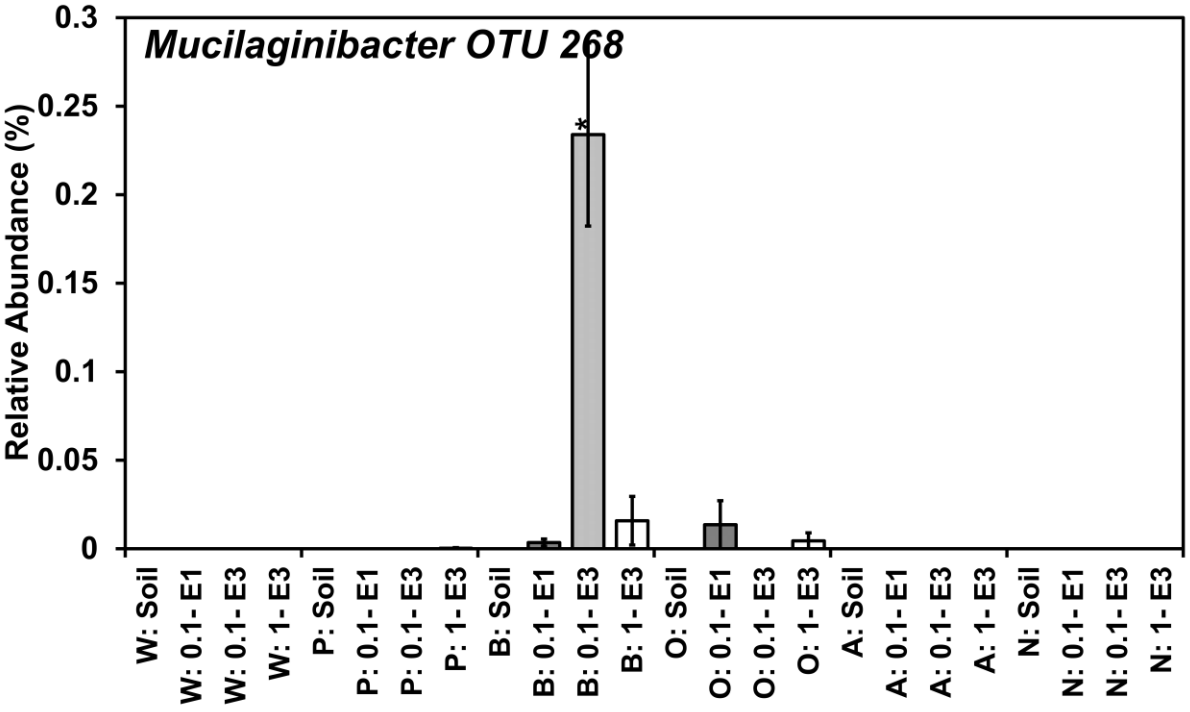


Figure 2.SI.97 *Mucilaginibacter* OTU 268 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first

enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Mucilaginibacter* OTU 268 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

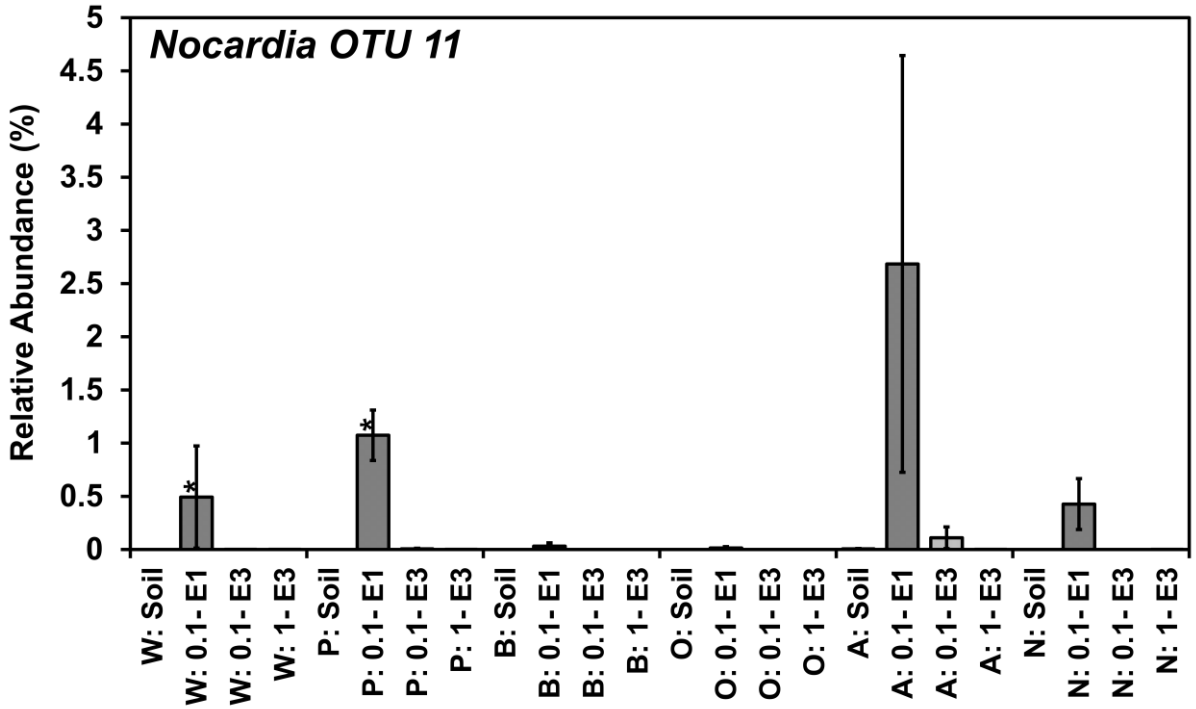


Figure 2.SI.98 *Nocardia* OTU 11 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Nocardia* OTU 11 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

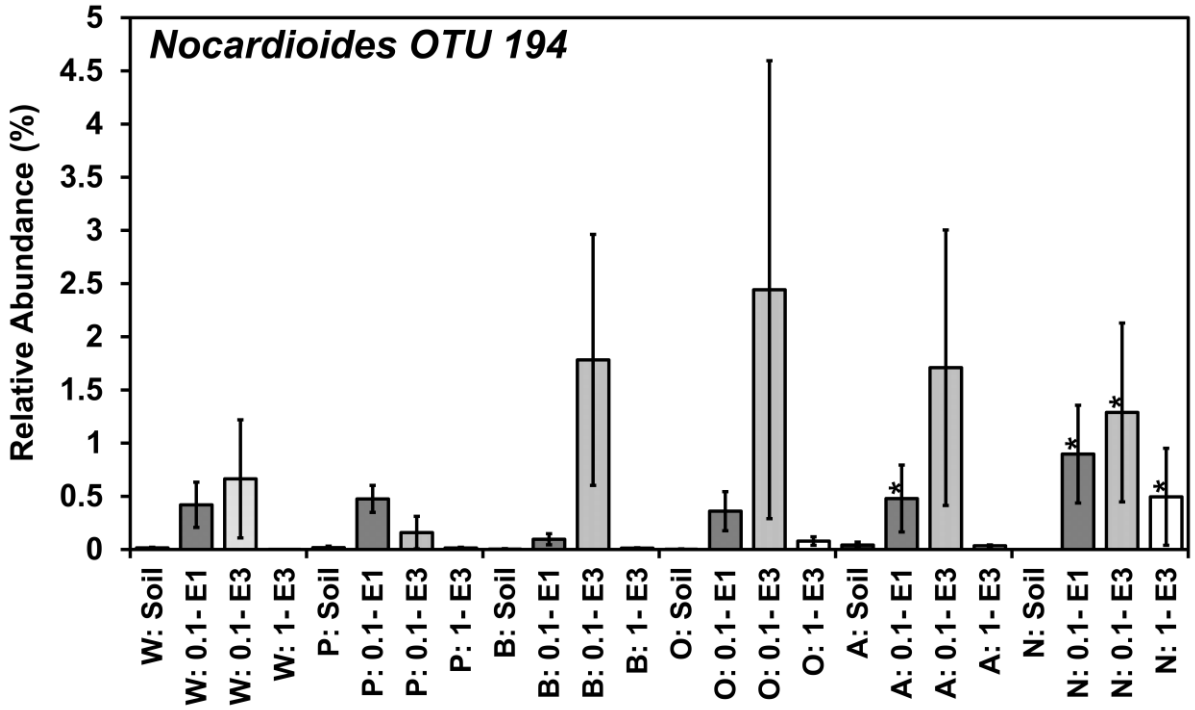


Figure 2.SI.99 *Nocardioides* OTU 194 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Nocardioides* OTU 194 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

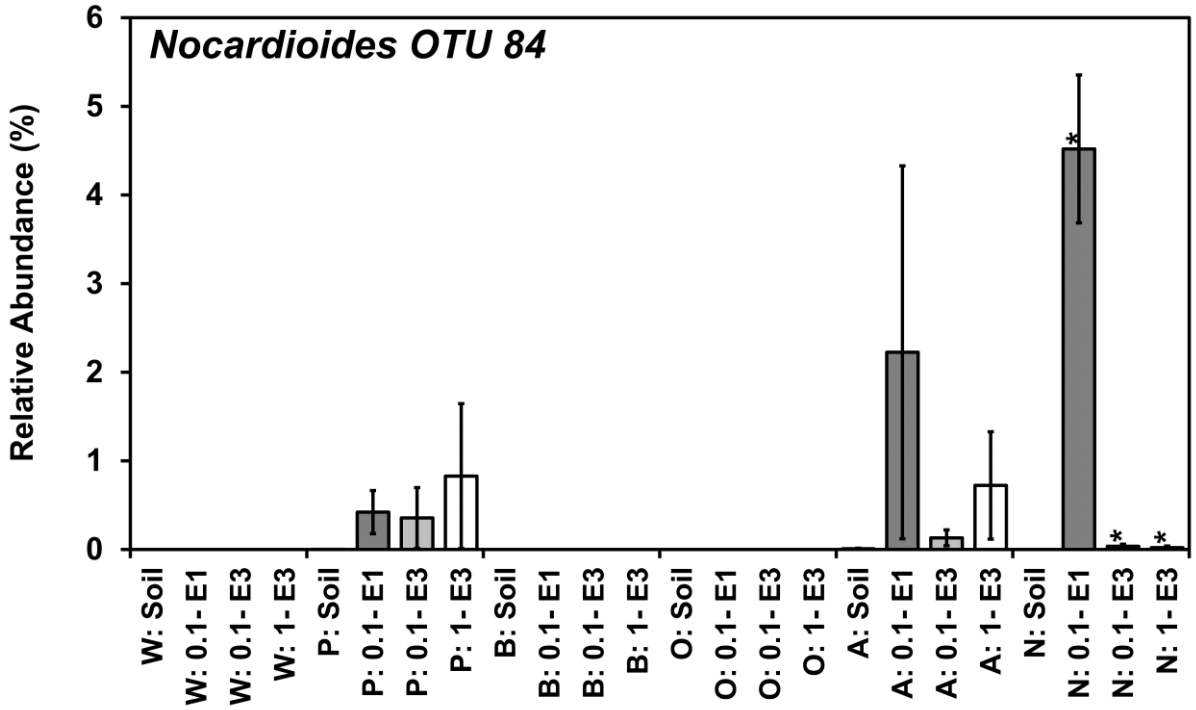


Figure 2.SI.100 *Nocardioides* OTU 84 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Nocardioides* OTU 84 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

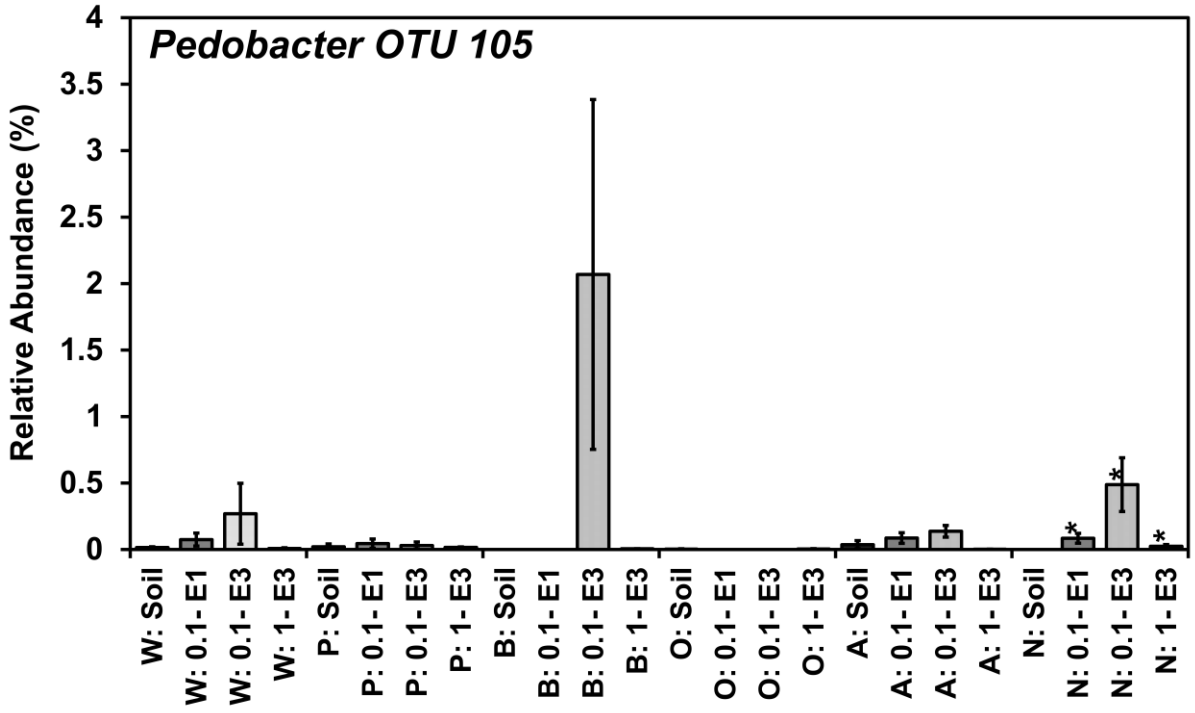


Figure 2.SI.101 *Pedobacter* OTU 105 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Pedobacter* OTU 105 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

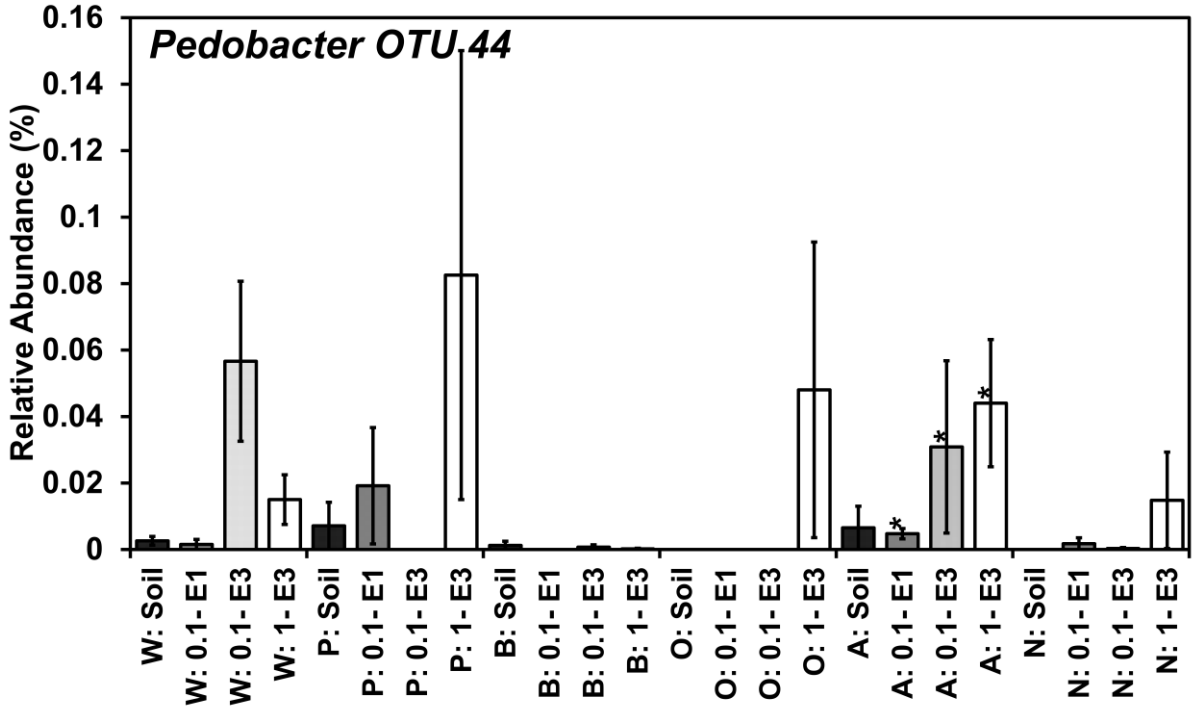


Figure 2.SI.102 *Pedobacter* OTU 44 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Pedobacter* OTU 44 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

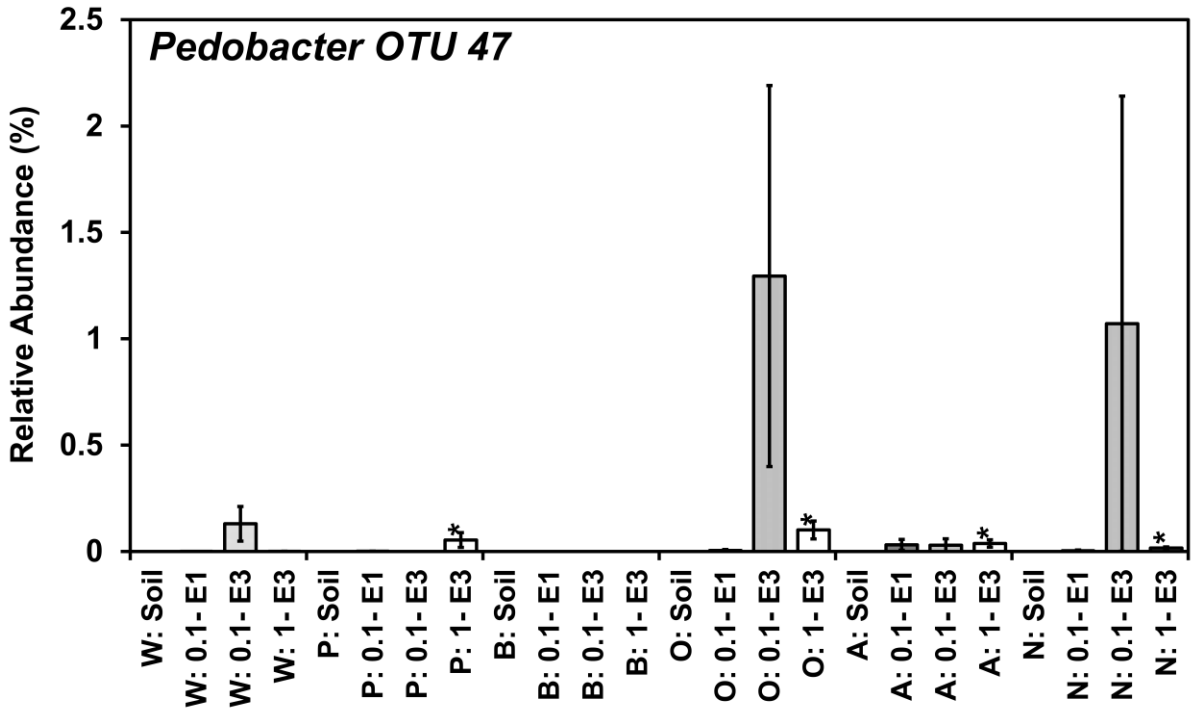


Figure 2.SI.103 *Pedobacter* OTU 47 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Pedobacter* OTU 47 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

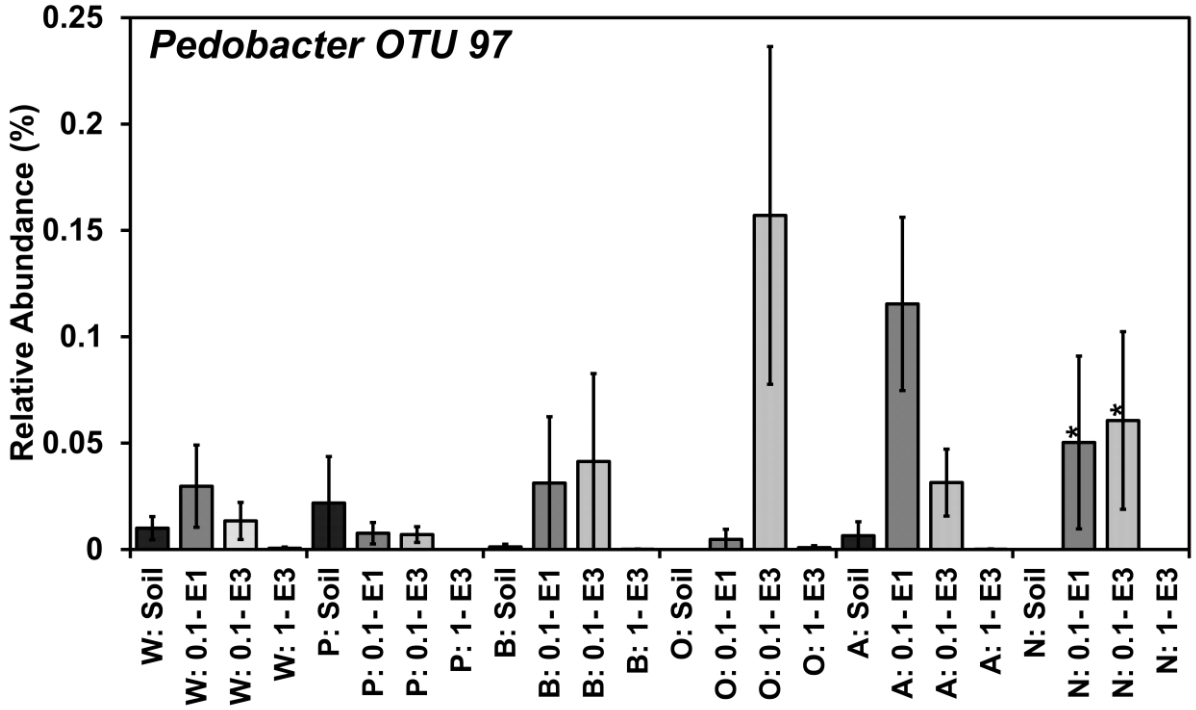


Figure 2.SI.104 *Pedobacter* OTU 97 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Pedobacter* OTU 97 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

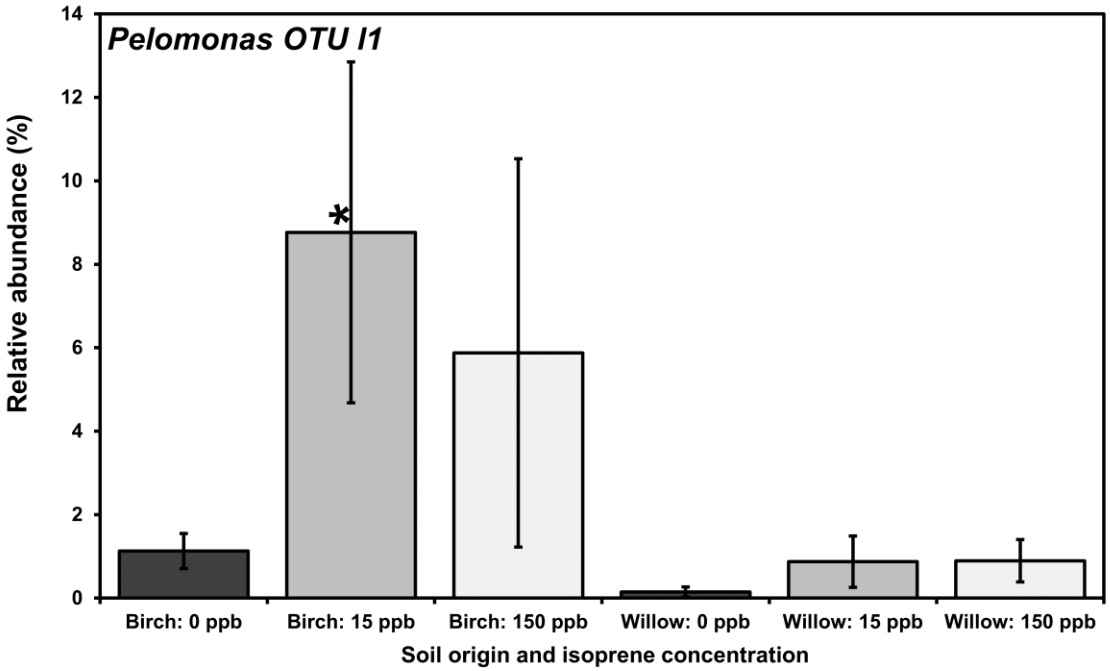


Figure 2.SI.105 *Pelomonas* OTU l1 relative abundances after incubation with 0, 15, and 150 ppb isoprene, replenished daily for three weeks, n=3 , Error bars = SE, *Pelomonas* OTU l1 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

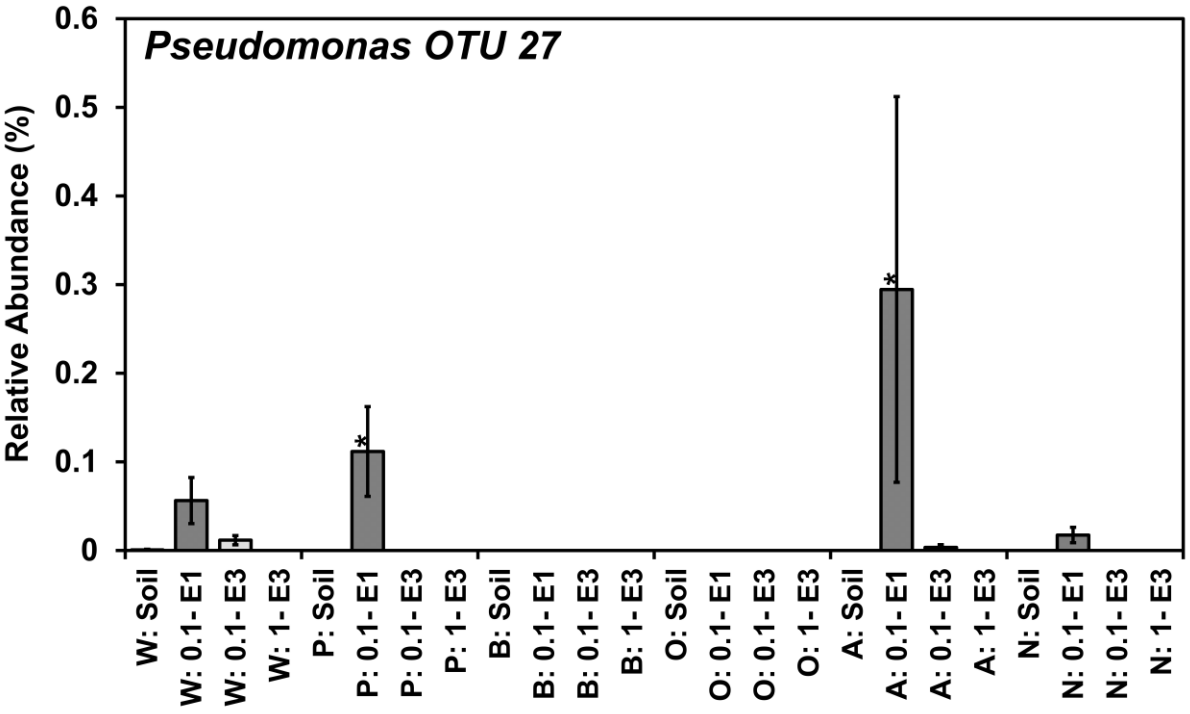


Figure 2.SI.106 *Pseudomonas* OTU 27 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated

isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Pseudomonas* OTU 27 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

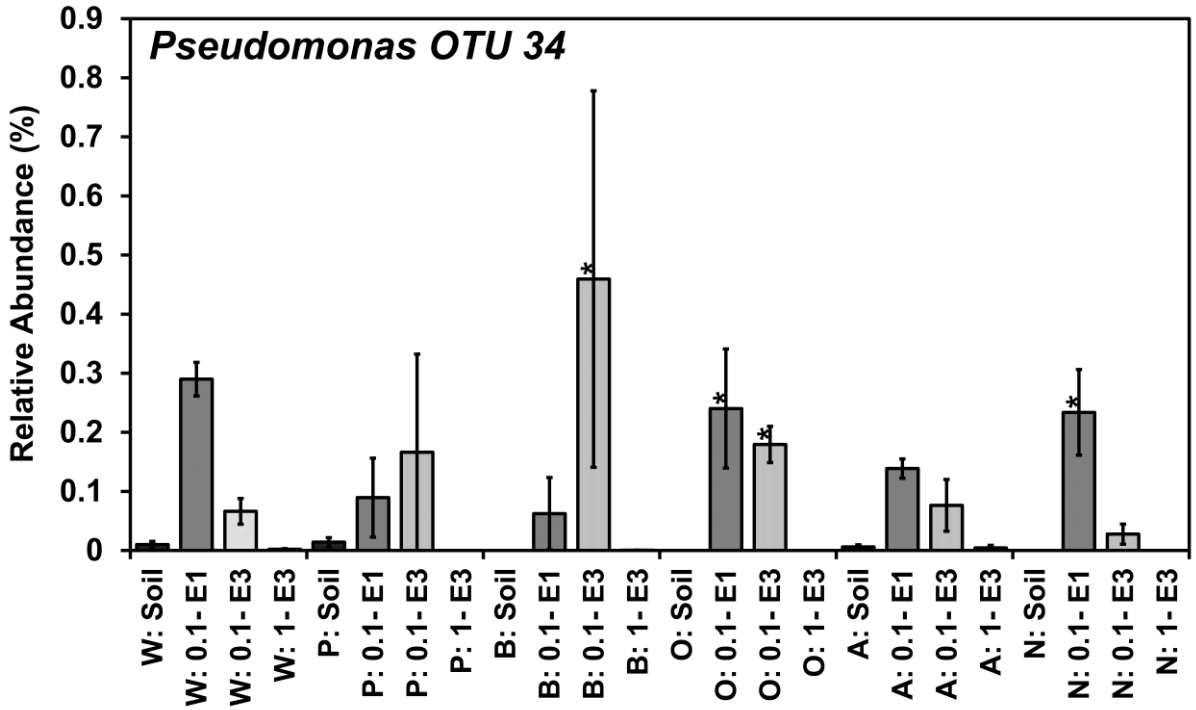


Figure 2.SI.107 *Pseudomonas* OTU 34 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Pseudomonas* OTU 34 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

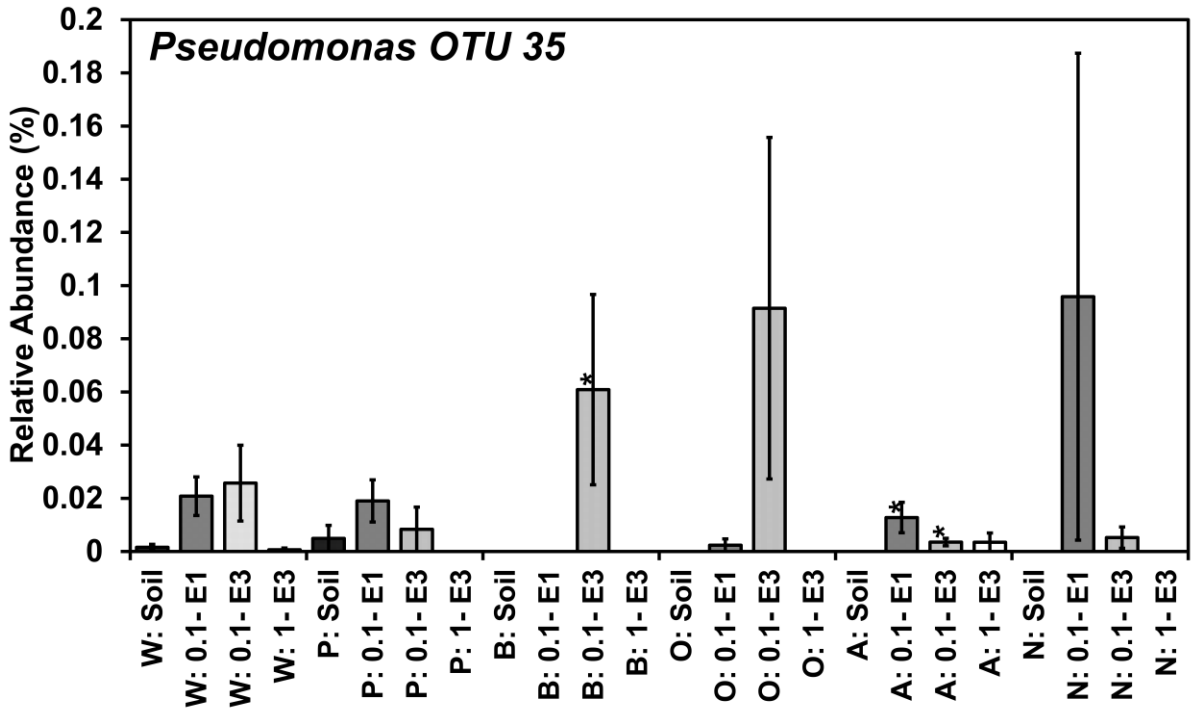


Figure 2.SI.108 *Pseudomonas* OTU 35 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Pseudomonas* OTU 35 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

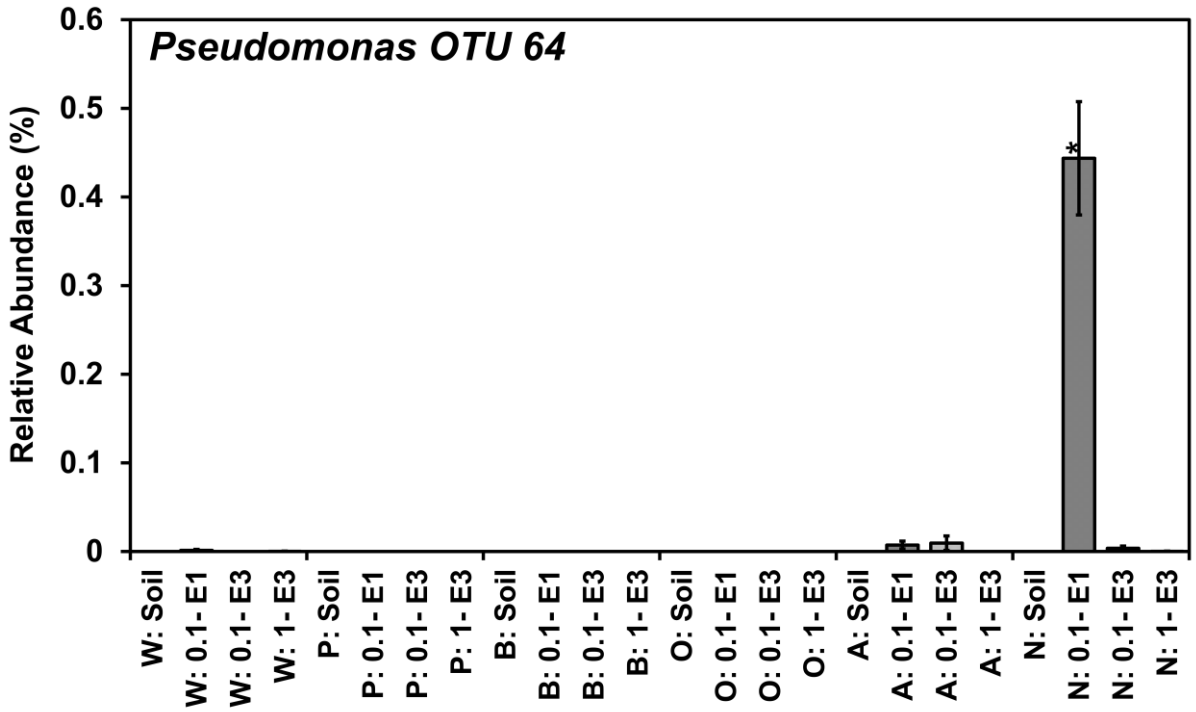


Figure 2.SI.109 *Pseudomonas* OTU 64 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Pseudomonas* OTU 64 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

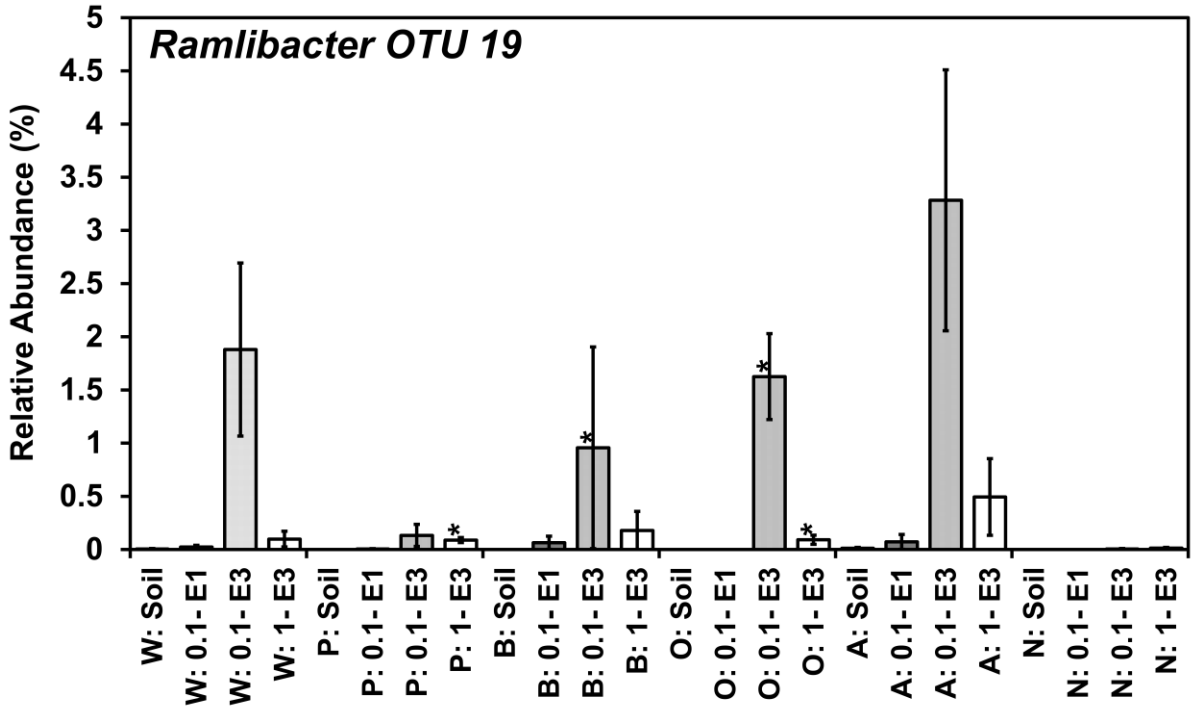


Figure 2.SI.140 *Ramlibacter* OTU 19 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Ramlibacter* OTU 19 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

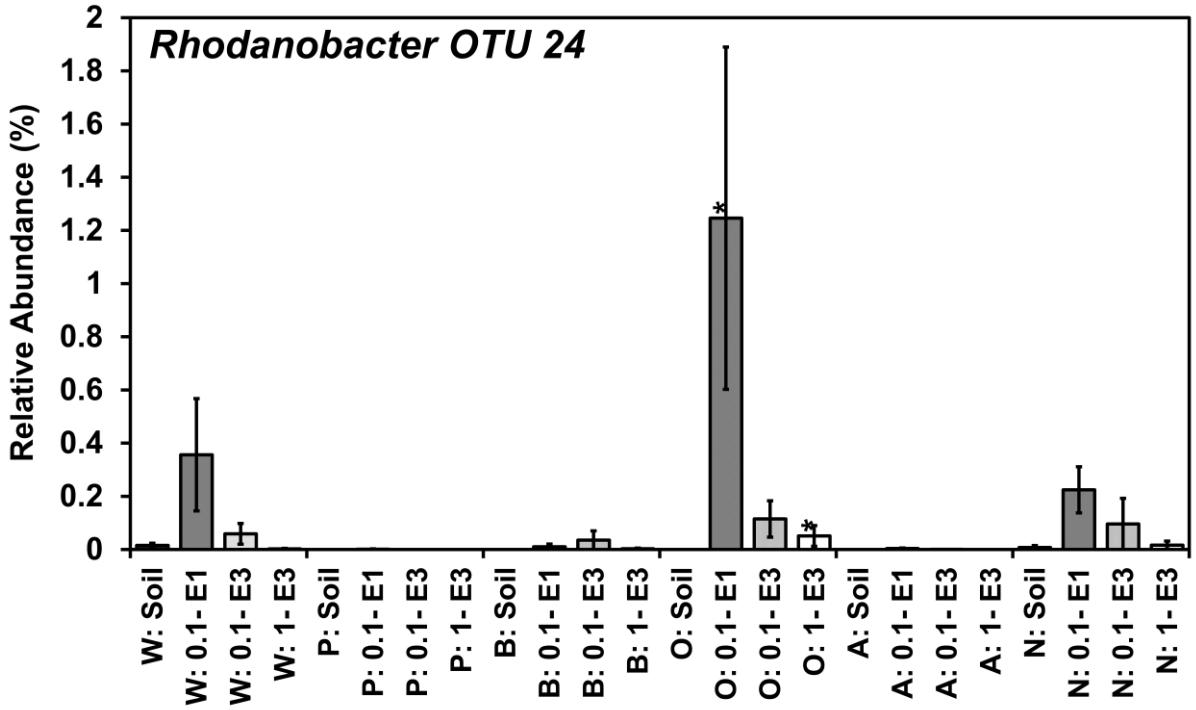


Figure 2.SI.141 *Rhodanobacter* OTU 24 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Rhodanobacter* OTU 24 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

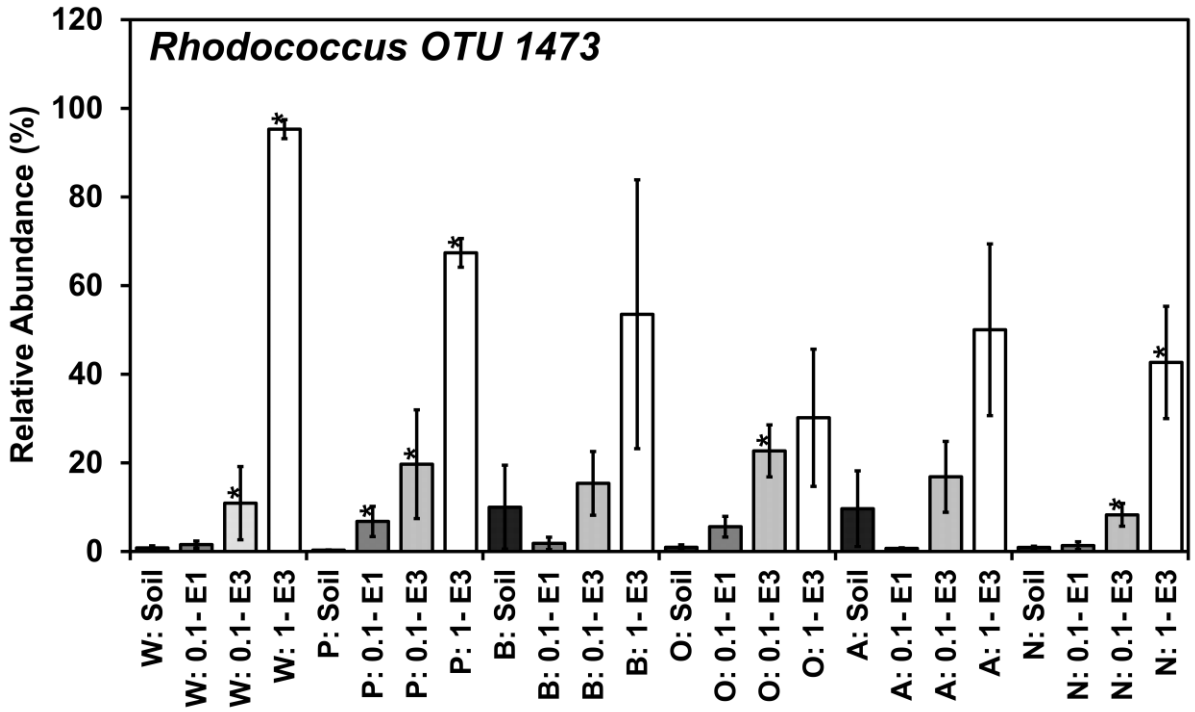


Figure 2.SI.142 *Rhodococcus* OTU 1473 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Rhodococcus* OTU 1473 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

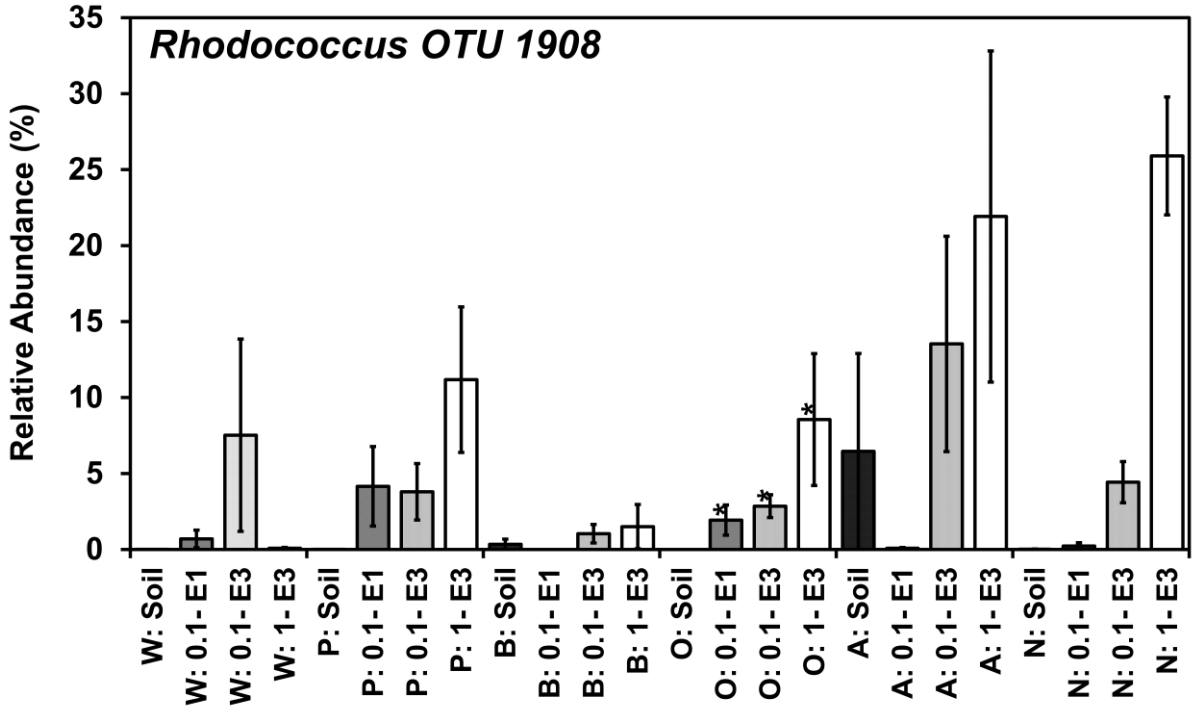
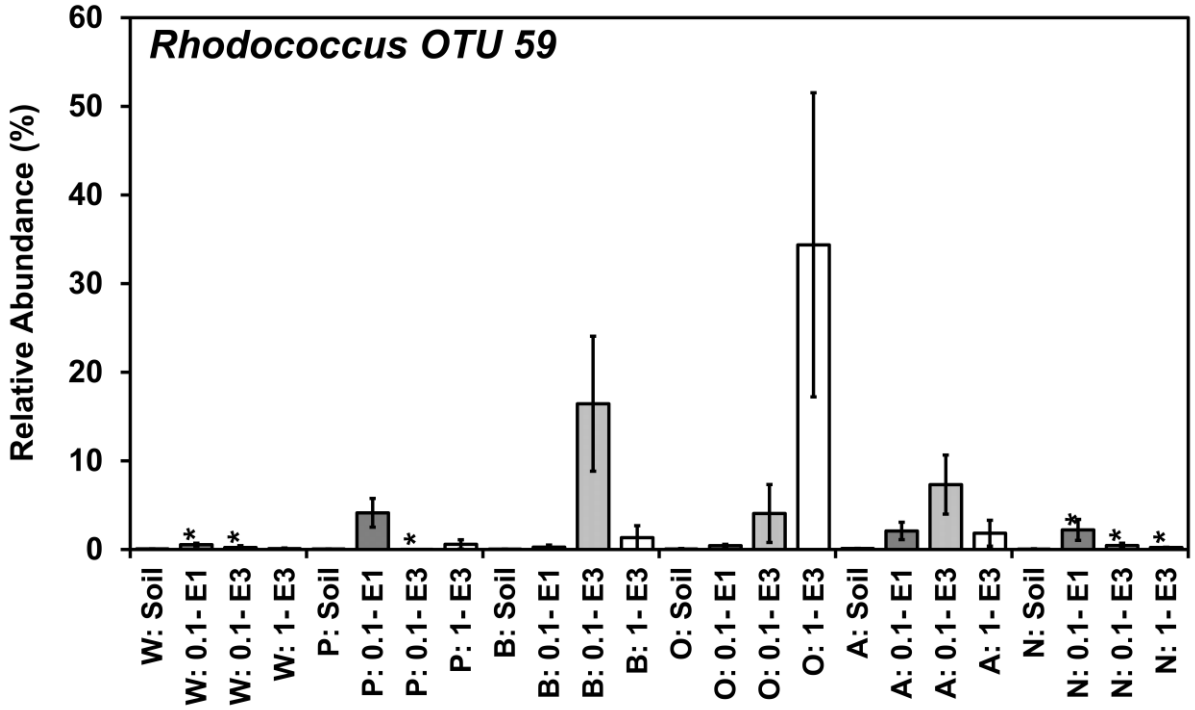


Figure 2.SI.143 *Rhodococcus* OTU 1908 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Rhodococcus* OTU 1908 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.



4Figure 2.SI.144 *Rhodococcus* OTU 59 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Rhodococcus* OTU 59 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

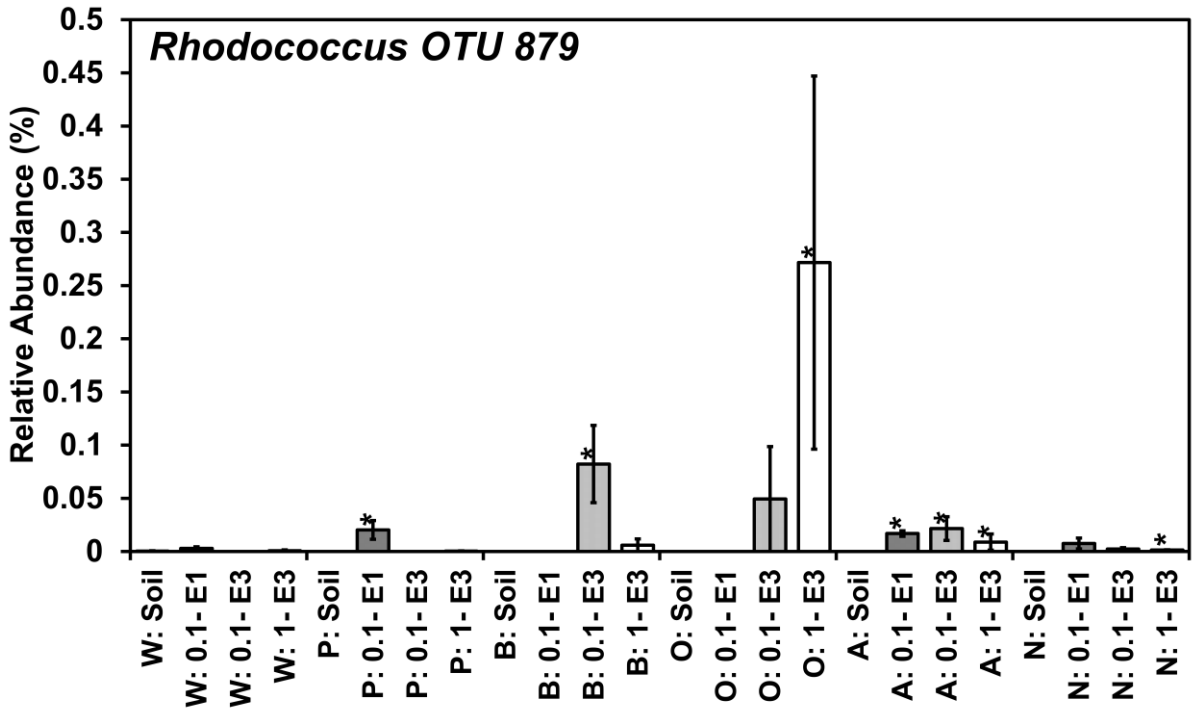
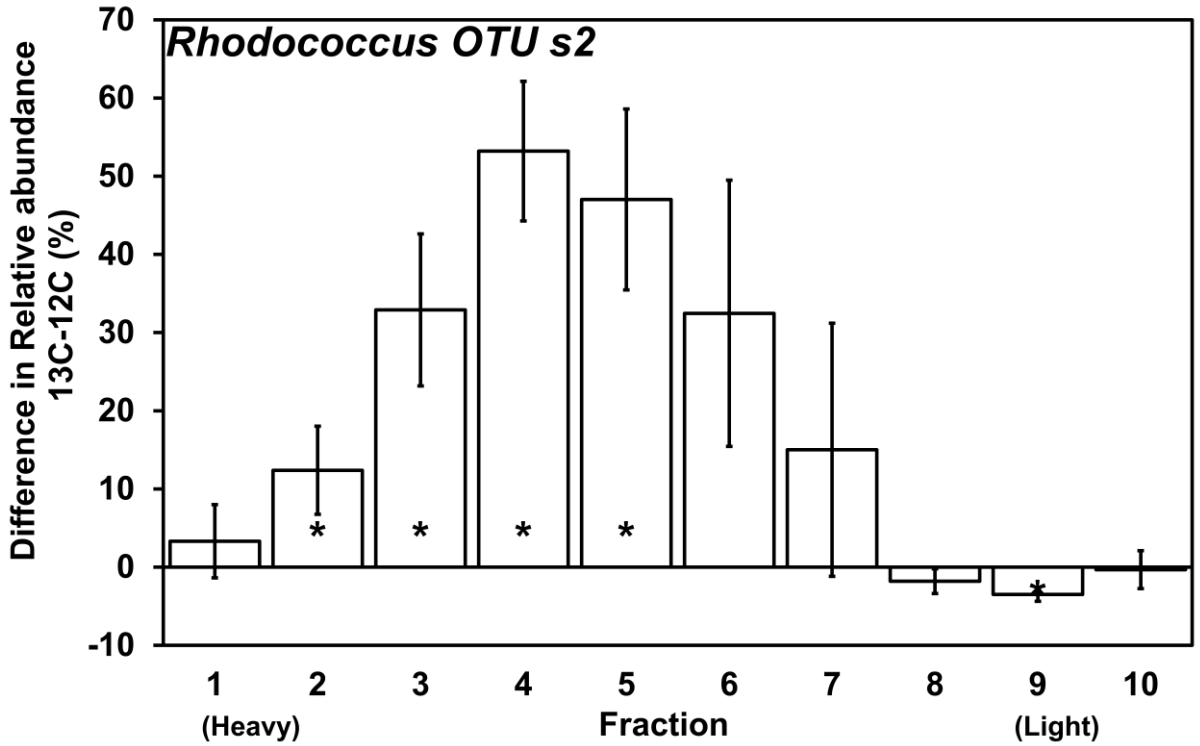


Figure 2.SI.145 *Rhodococcus* OTU 879 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Rhodococcus* OTU 879 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.



4Figure 2.SI.146 *Rhodococcus* OTU s2 relative abundances in 13C fractions net of *Rhodococcus* OTU s1 relative abundances in corresponding 12C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, 13C and 12C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Rhodococcus* OTU s1 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

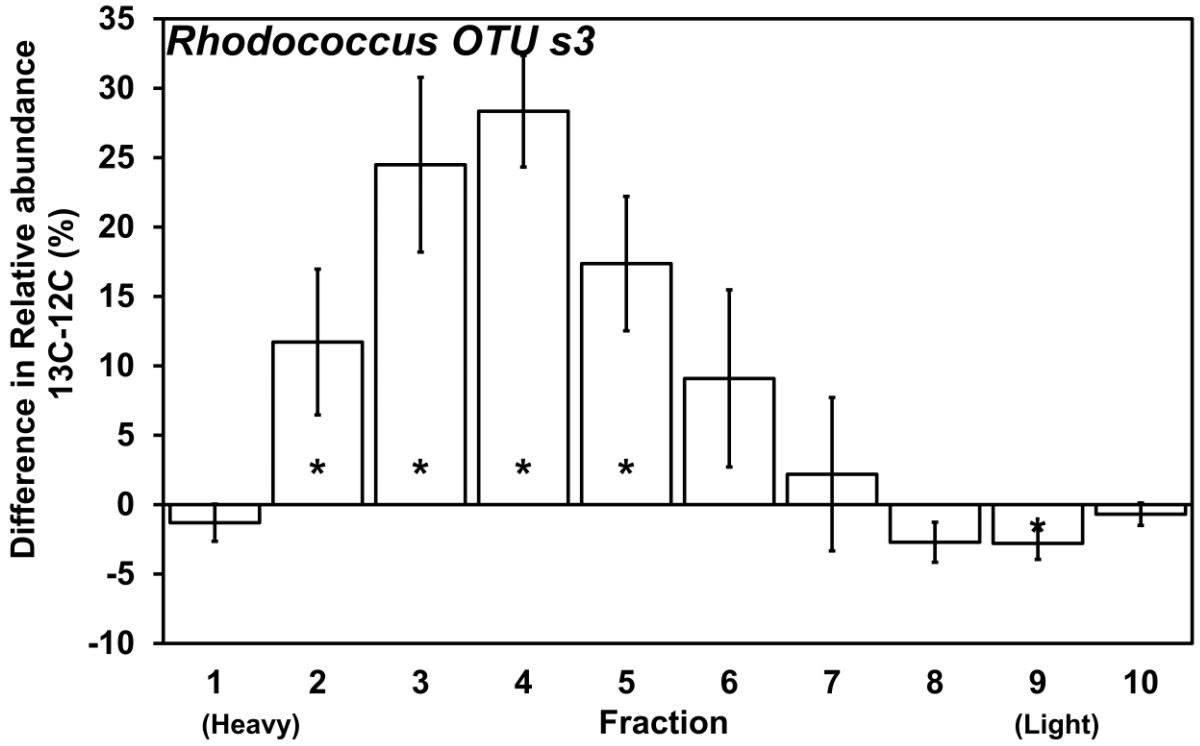


Figure 2.SI.147 *Rhodococcus* OTU s3 relative abundances in 13C fractions net of *Rhodococcus* OTU s2 relative abundances in corresponding 12C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, 13C and 12C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Rhodococcus* OTU s2 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

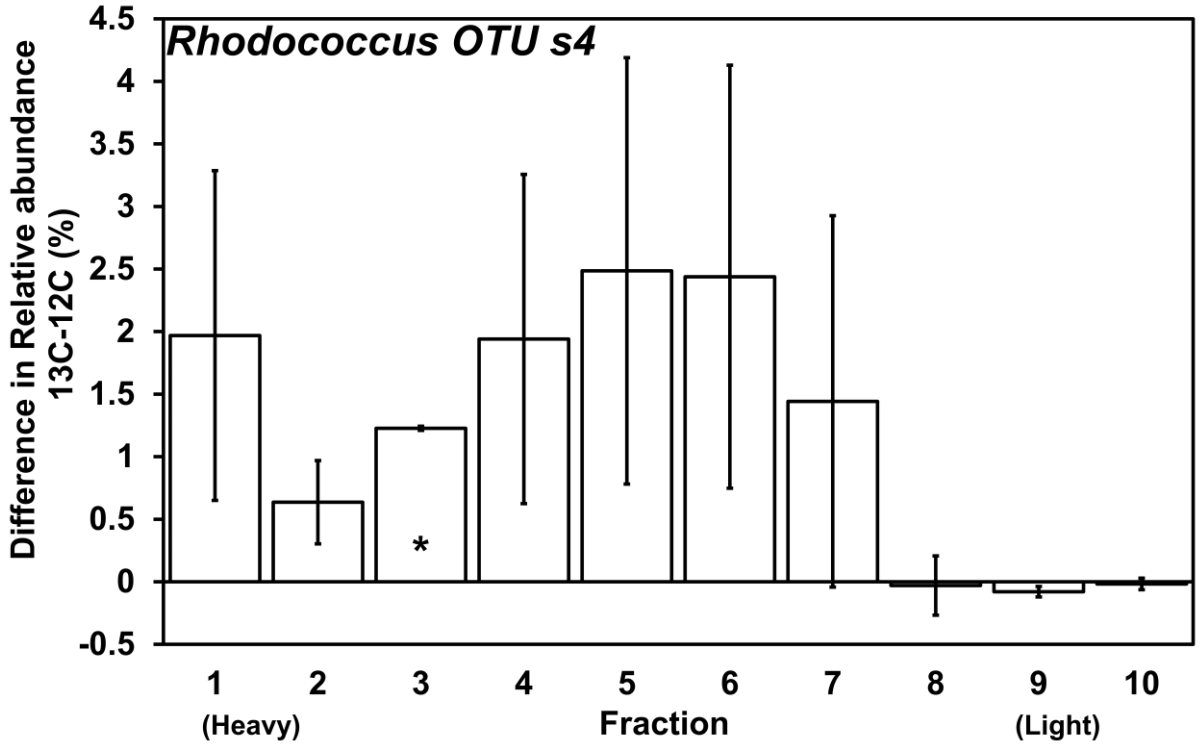


Figure 2.SI.148 *Rhodococcus* OTU s4 relative abundances in 13C fractions net of *Rhodococcus* OTU s3 relative abundances in corresponding 12C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, 13C and 12C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Rhodococcus* OTU s3 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

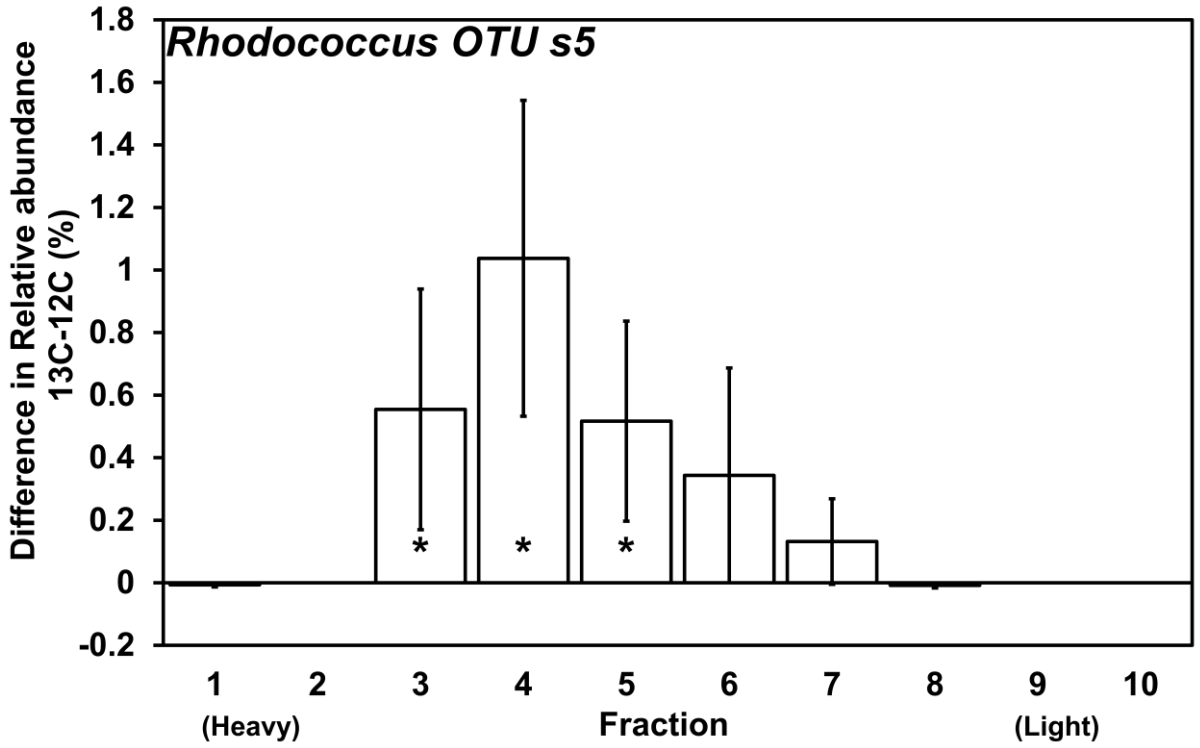


Figure 2.SI.149 *Rhodococcus* OTU s5 relative abundances in 13C fractions net of *Rhodococcus* OTU s4 relative abundances in corresponding 12C fractions, after separate enrichment with 1 ml 30°C saturated isoprene headspace for 4 days, 13C and 12C isoprene and density gradient centrifugation, n = 3, Error bars = SE, *Rhodococcus* OTU s4 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

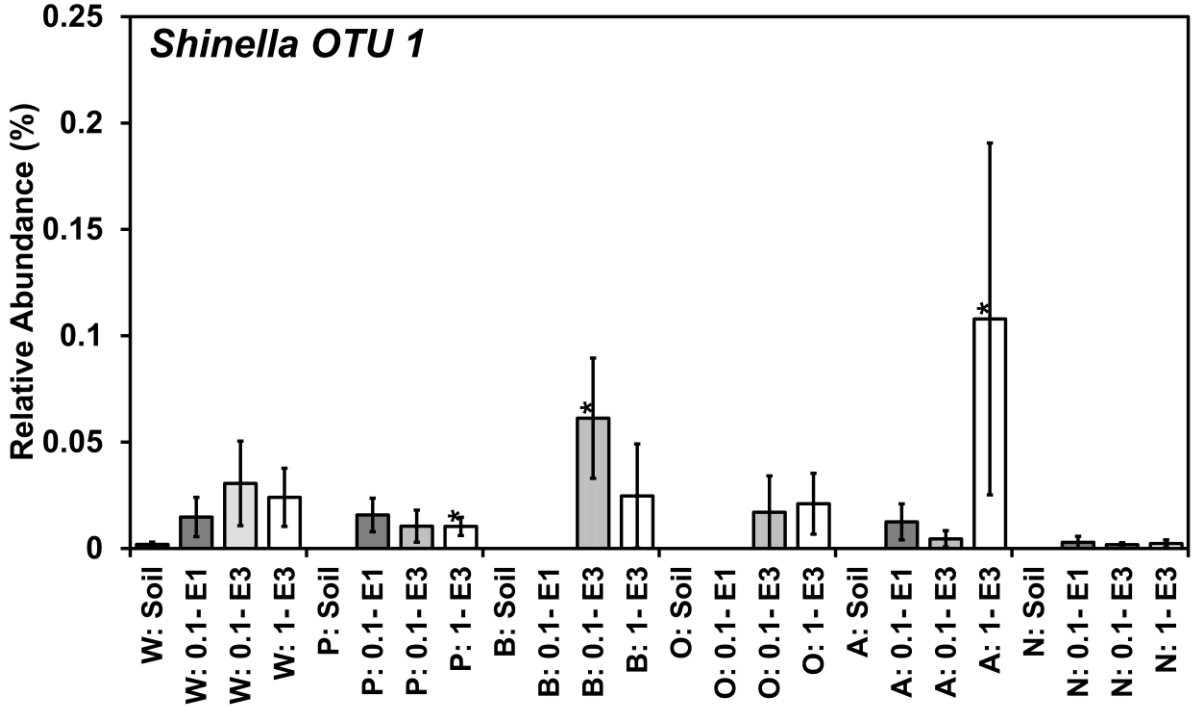


Figure 2.SI.150 *Shinella* OTU 1 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Shinella* OTU 1 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

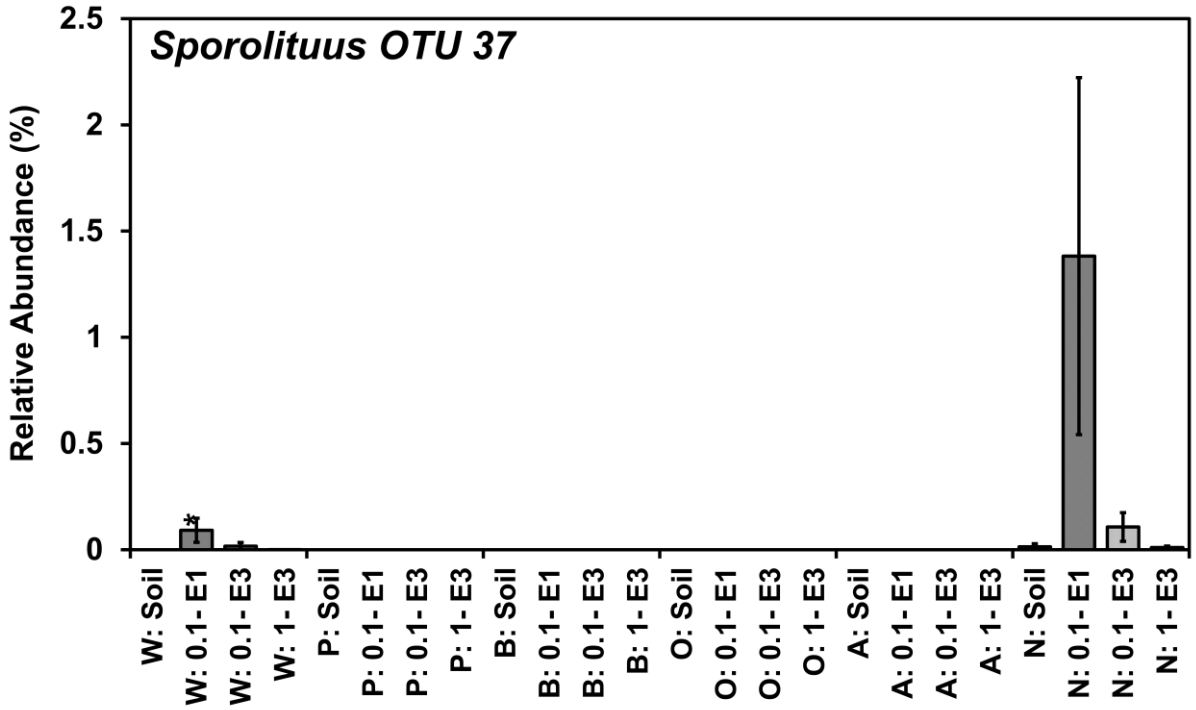


Figure 2.SI.151 *Sporolituus* OTU 37 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Sporolituus* OTU 37 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

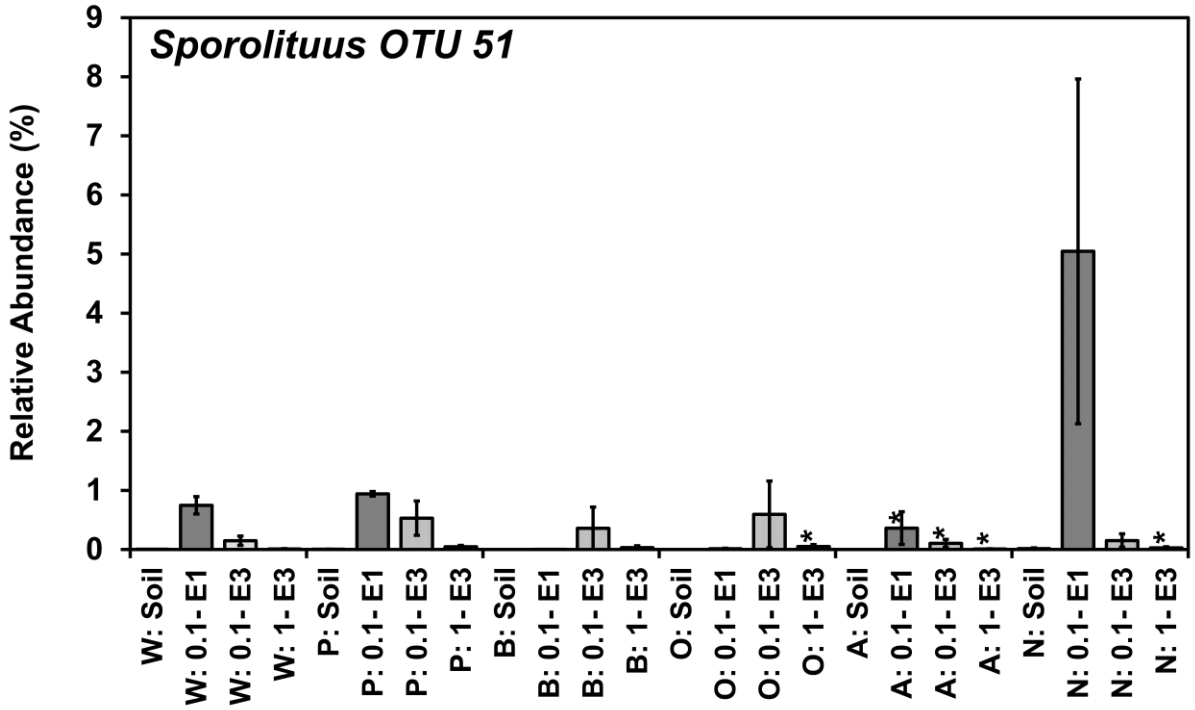
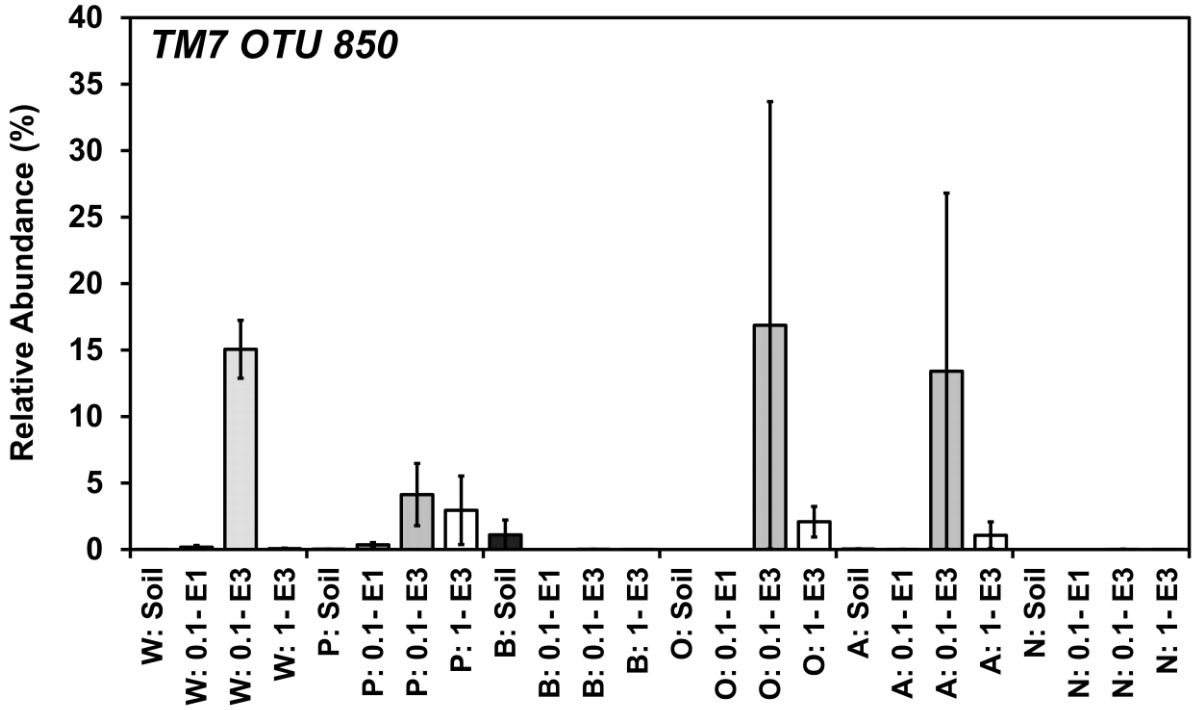


Figure 2.SI.152 *Sporolituus* OTU 51 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Sporolituus* OTU 51 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.



15Figure 2.SI.153 TM7 OTU 850 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, TM7 OTU 850 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

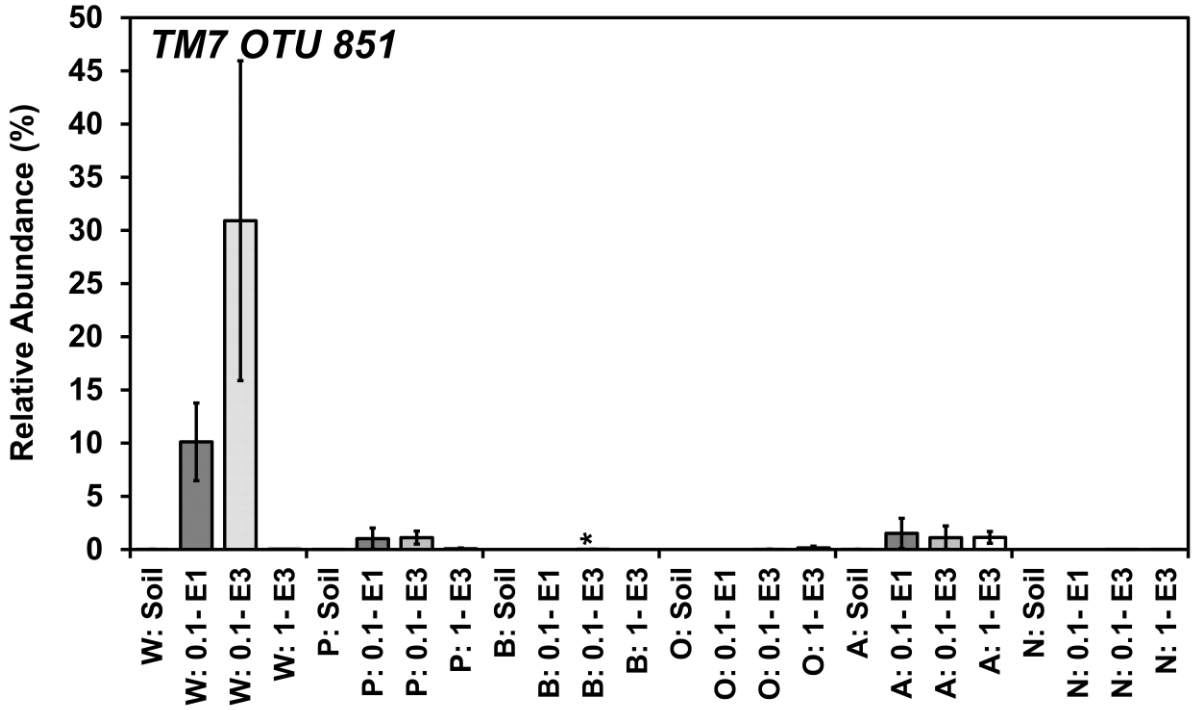


Figure 2.SI.154 TM7 OTU 851 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, TM7 OTU 851 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

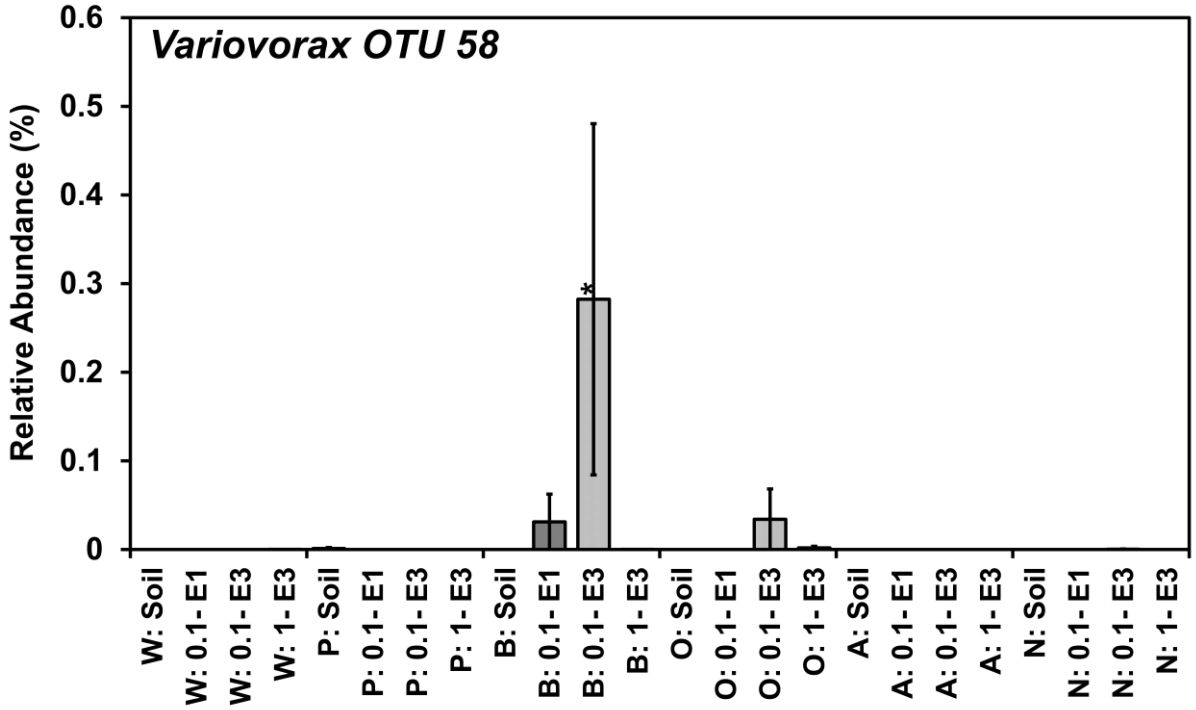


Figure 2.SI.155 *Variovorax* OTU 58 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Variovorax* OTU 58 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

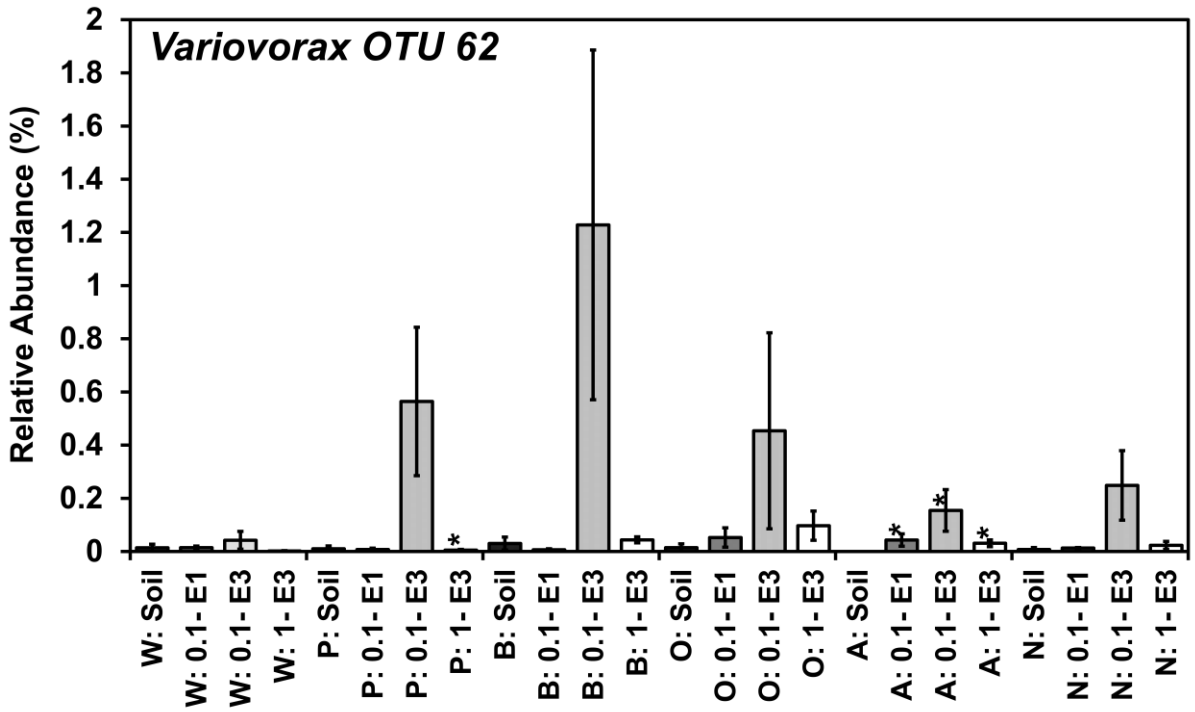


Figure 2.SI.156 *Variovorax* OTU 62 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Variovorax* OTU 62 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

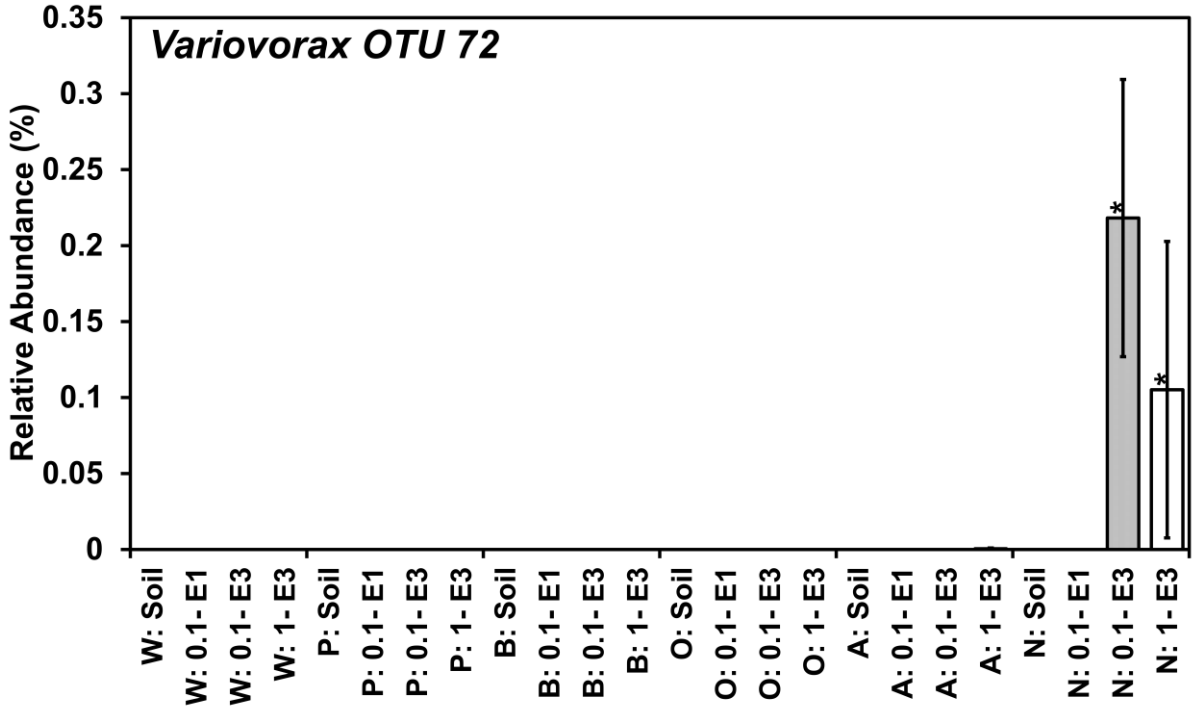


Figure 2.SI.157 *Variovorax* OTU 72 relative abundances in sequential enrichment of soil with isoprene. Soil was obtained from under the canopy of trees W = Willow, P = Poplar, B = Birch, O = Oak, A = Ash, N = No-Trees. Enrichment was at two levels 1 ml (0.8%) 30°C saturated isoprene headspace addition = 1.0, 0.1 ml (0.08%) 30°C saturated isoprene headspace addition = 0.1. Pre-enriched soil abundance = soil, first enrichment = E1, Second enrichment = E2, Third enrichment = E3. * = significance at p<0.05 (Kruskal-Wallis), n = 3, Error bars = SE, *Variovorax* OTU 72 defined by Swarm clustering (d=1) and RDP classification at deepest assignment.

Table 2.SI.1 Soil charecteristics for no-trees sampling area on the four dates of sampling throughout 2015, and the weather data for the sampling date and preceeding ten days

Date of sampling	Soil pH (mean)	Porosity (%)	Soil Temperature °C	Weather data for sampling day and the preceding 10 days											
24/03/2015	4.89	66	19	**Date**	14/03/2015	15/03/2015	16/03/2015	17/03/2015	18/03/2015	19/03/2015	20/03/2015	21/03/2015	22/03/2015	23/03/2015	24/03/2015
				Precipitation (mm)	0	0	0	1.02	0	0	0	0.25	0	0.25	0.76
				Air Temp (mean) °C	4	6	6	6	5	6	7	5	3	4	6
18/06/2015	4.89	70	21	**Date**	08/06/2015	09/06/2015	10/06/2015	11/06/2015	12/06/2015	13/06/2015	14/06/2015	15/06/2015	16/06/2015	17/06/2015	18/06/2015
				Precipitation (mm)	0	4.06	0	2.03	0	0	0	0	0	1.02	0
				Air Temp (mean) °C	17	17	19	17	14	23	23	27	23	19	14
29/09/2015	4.86	69	20	**Date**	19/09/2015	20/09/2015	21/09/2015	22/09/2015	23/09/2015	24/09/2015	25/09/2015	26/09/2015	27/09/2015	28/09/2015	29/09/2015
				Precipitation (mm)	0	0.25	6.1	0.51	0	3.05	0	0	0	0	0
				Air Temp (mean) °C	13	13	12	10	12	12	11	11	11	12	13
10/12/2015	4.89	62	9	**Date**	30/11/2015	01/12/2015	02/12/2015	03/12/2015	04/12/2015	05/12/2015	06/12/2015	07/12/2015	08/12/2015	09/12/2015	10/12/2015
				Precipitation (mm)	5.08	1.02	0	1.02	0	0	0	0	2.03	0	0.76
				Air Temp (mean) °C	8	9	11	11	8	10	12	11	9	6	9

Measurement of pH was performed in $CaCl_2$ solution. Weather data was obtained from historical weather data based on data obtained from the closest weather station (Wattisham, 13 miles north of sampling location)

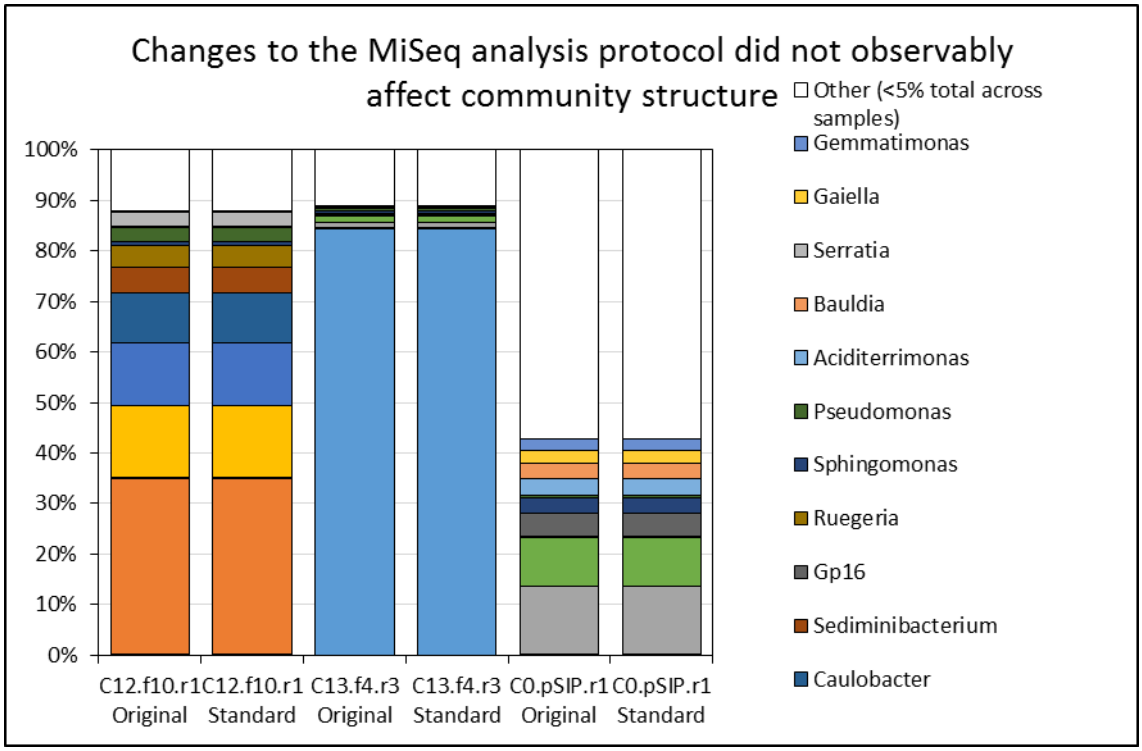


Figure 2.SI.158 Changes to the protcol for in this thesis from that suggested by Schirmer et al., (2015), namely the use of Bayes Hammer as a pre-processing step before Sickle (instead of after Sickle), did not have any observable effect on community structure (Original = approach taken in this thesis, Standard = Schirmer approach).

3.SI: Chapter 3 Supplementary Information

Table 3.SI.2 Carbon source utilisation by selected isolates. Isolates were incubated with a range of carbon sources using Biolog carbon-source testing plates PM1 and PM2.

Substrate	1(2b)b	bl28a	bl28ba	GM3	W9	Tc13b
Capric Acid	**	**	**		**	**
2-Hydroxy Benzoic Acid	*	**	**		**	**
2-Deoxy Adenosine	**		**		**	**
2-Deoxy-DRibose	*	*	*			**
Sorbic Acid	*		*		*	**
2,3-Butanone	*		**		*	**
D-Galactonic Acid-g-Lactone	**				**	*
D-Ribose				*	**	*
Tween 80			*		*	*
m-Hydroxy Phenyl Acetic Acid		*	*		**	
Dihydroxy Acetone	*		**			*
L-Arabinose					**	*
D-Trehalose			*		**	
D-Xylose				*	**	
Tween 20					**	*
D-Fructose		*	**			
L-Malic Acid		*	*			
L-Lyxose				*	**	
b-Cyclodextrin		*	*			
Pectin			*			*
3-0-b-D-Galactopyranosyl-					*	**
D-Glucosamine	*					*
L-Aspartic Acid					**	
L-Proline			*			
D-Alanine					*	
D-Mannose			*			
D-Serine					**	
Glycerol						**
D-Gluconic Acid			*			
D,L-a-GlycerolPhosphate					*	
Formic Acid					**	
D-Mannitol			*			
L-Glutamic Acid					*	
DL-Malic Acid			*			
Acetic Acid						*
a-D-Glucose			*			
Maltose					*	

Substrate	1(2b)b	bl28a	bl28ba	GM3	W9	Tc13b
Tween 40						**
Lactulose					*	
Sucrose			*			
D-Fructose-6- Phosphate					*	
a-Hydroxy Butyric Acid					*	
D-Threonine					**	
Fumaric Acid			*			
Bromo Succinic Acid			*			
Glyoxylic Acid					**	
Acetoacetic Acid					**	
Mono Methyl Succinate			*			
Methyl Pyruvate					**	
Tyramine					**	
D-Galacturonic Acid					**	
Phenylethylamine					**	
g-Cyclodextrin					*	
Laminarin						*
N-Acetyl-DGalactosamine					*	
Gentiobiose					**	
Lactitol					*	
D-Melezitose		*				
Maltitol					*	
D-Raffinose					*	
Xylitol					*	
Citramalic Acid		*				
b-Hydroxy Butyric Acid					*	
Malonic Acid					*	
D-Ribono-1,4- Lactone			*			
Sebacic Acid		*				
Putrescine					*	

*** = p <0.05, ** = p < 0.01, *** = p < 0.001 compared to no carbon-source controls. Based on tetrazolium dye as an indicator of respiration. A complete list of carbon sources the selected isolates were screened on is available in the Appendix 3.AX.2.**

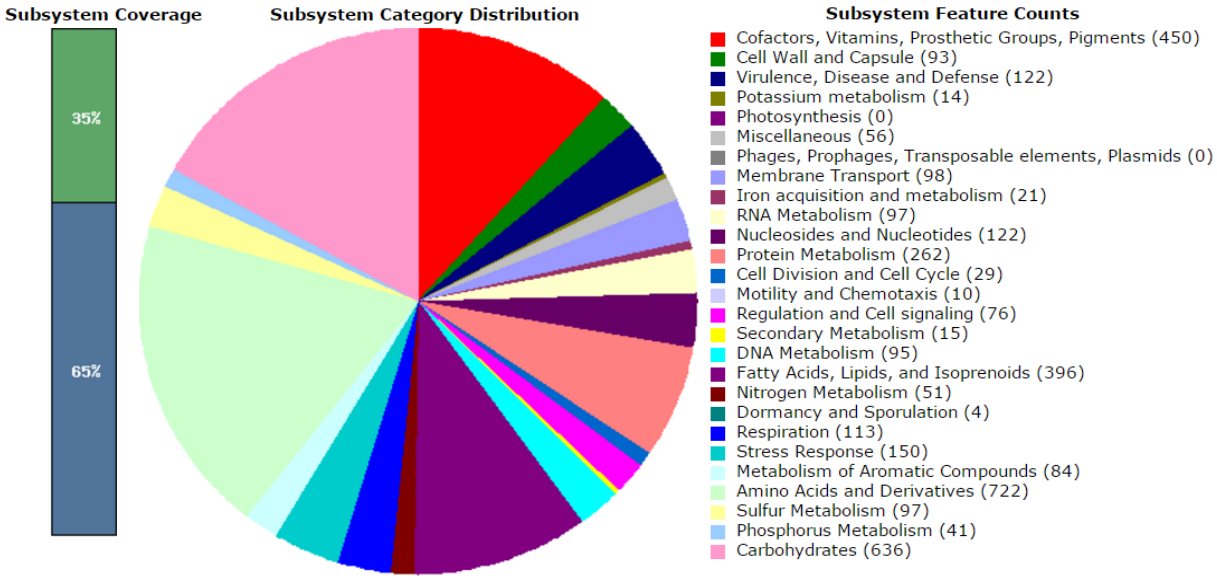


Figure 3.SI.1 RAST annotated protein predicted subsystem coverage, distribution and feature counts for bl28ba.

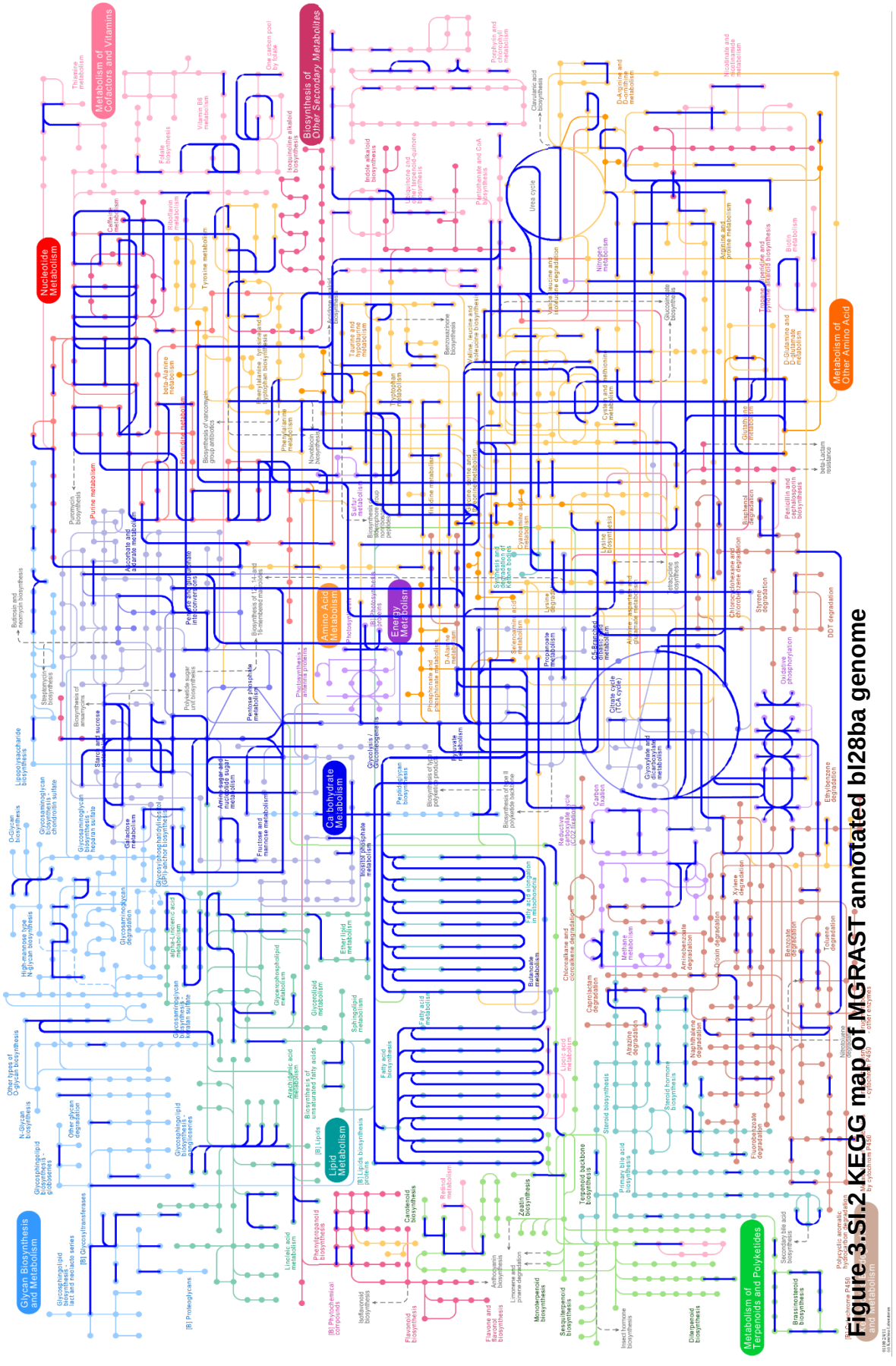


Figure 3.SI.2 KEGG map of MGRAST annotated bl28ba genome

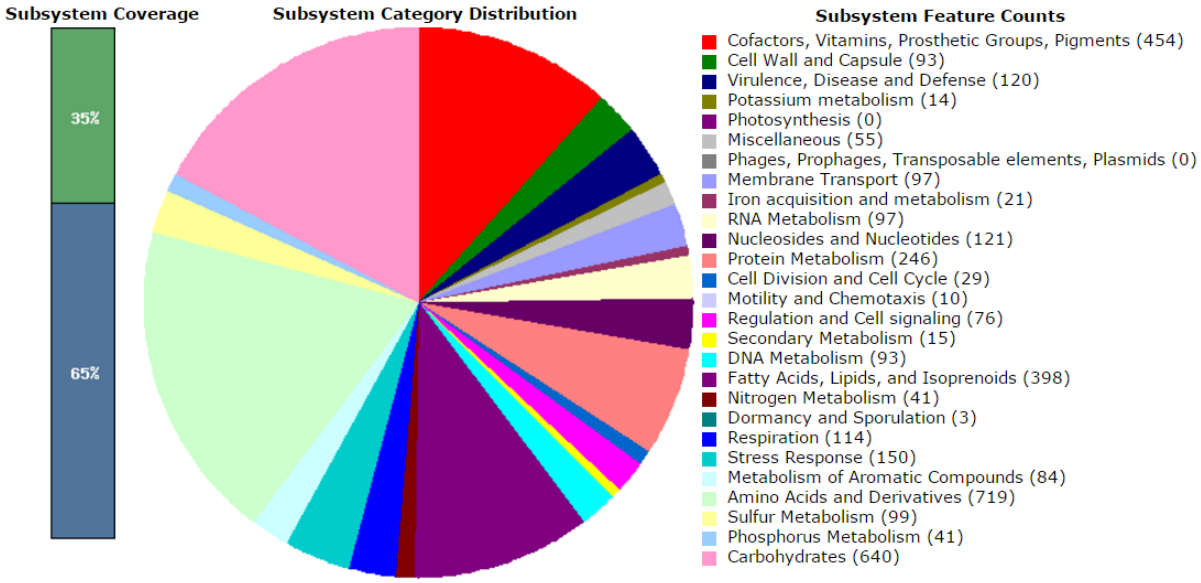


Figure 3.SI.3 RAST annotated protein predicted subsystem coverage, distribution and feature counts for bl28a.

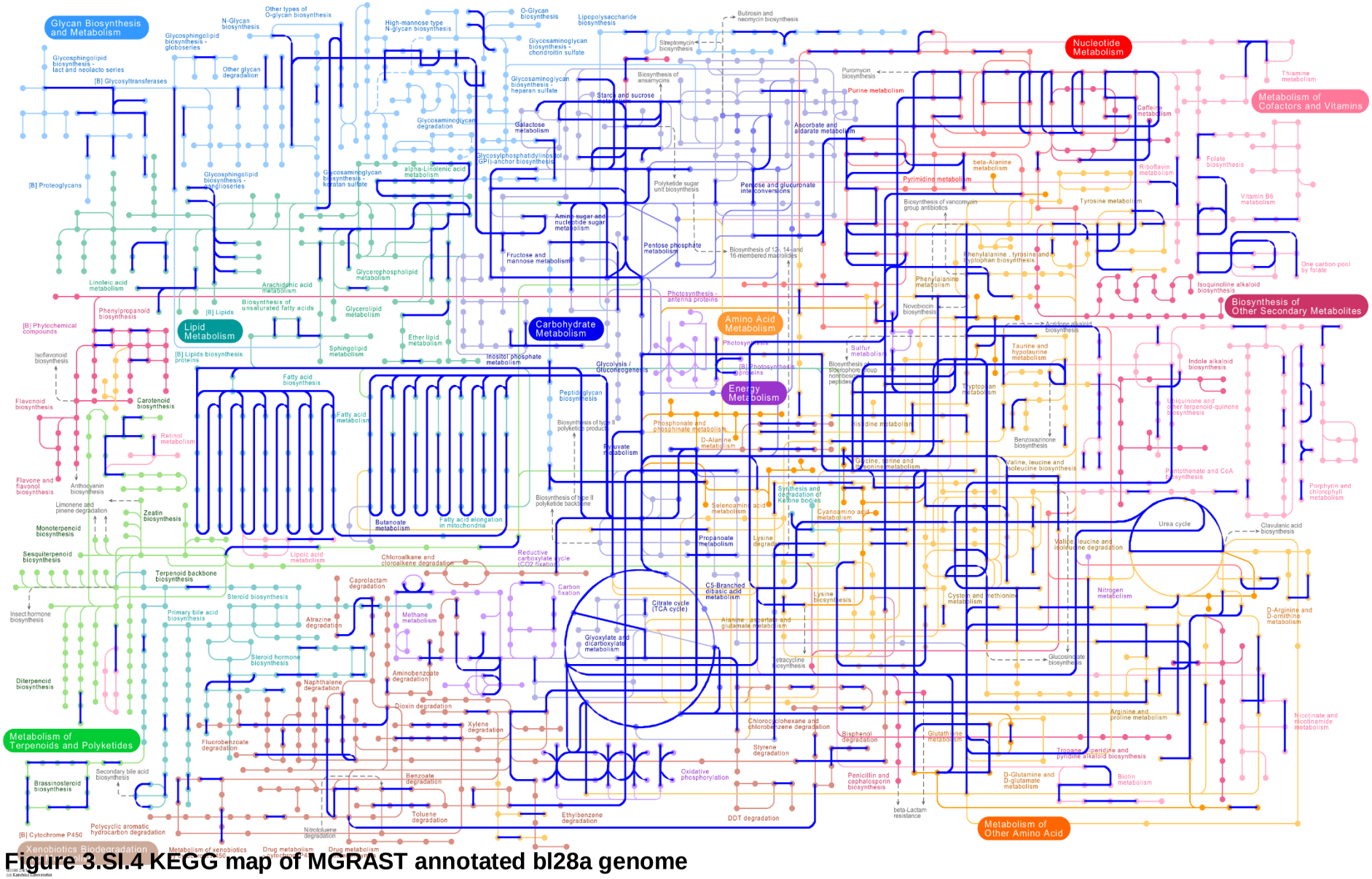


Figure 3.SI.4 KEGG map of MGRAST annotated bl28a genome

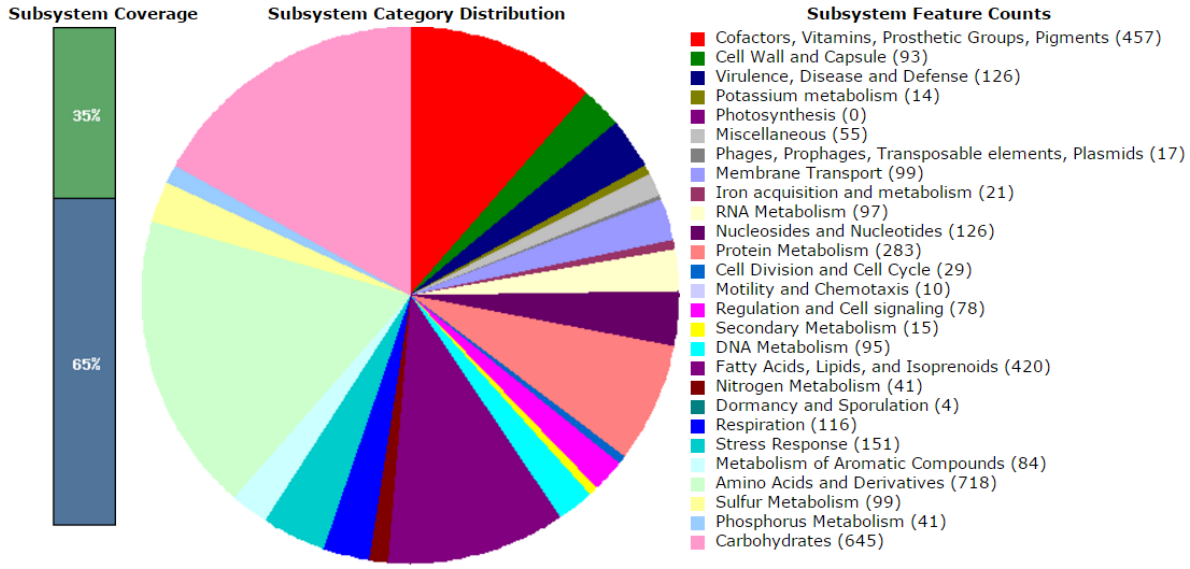


Figure 3.5 RAST annotated protein predicted subsystem coverage, distribution and feature counts for 1(2b)b.

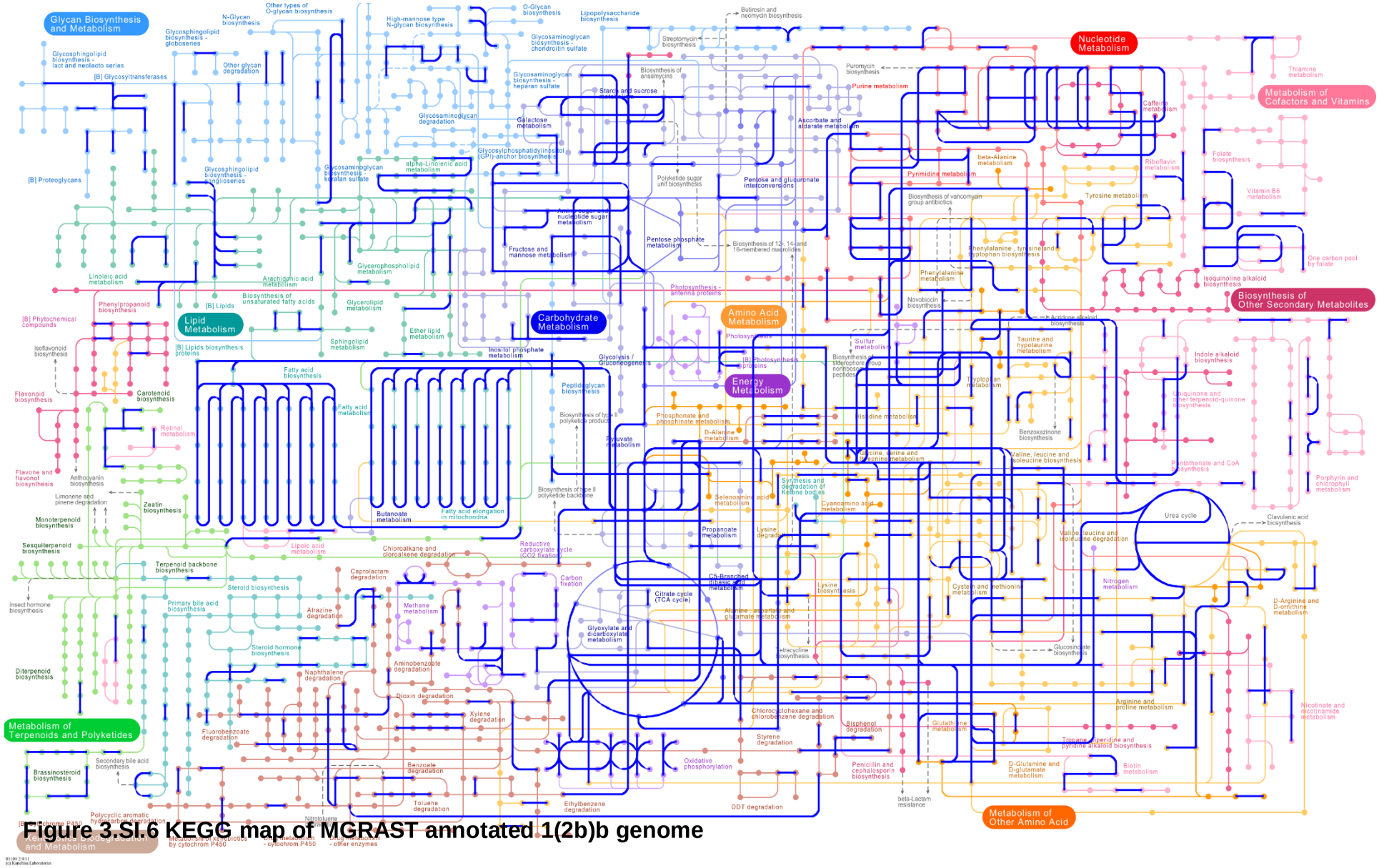


Figure 3.SI.6 KEGG map of MGRAST annotated 1(2b)b genome

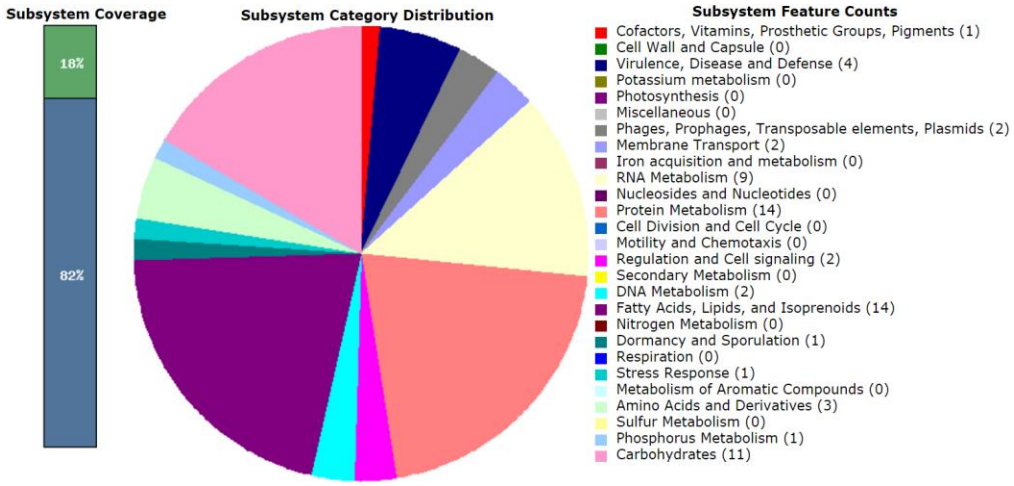


Figure 3.SI.7 RAST annotated protein predicted subsystem coverage, distribution and feature counts for TM7 stripped out of a TM7-rich microcosm.

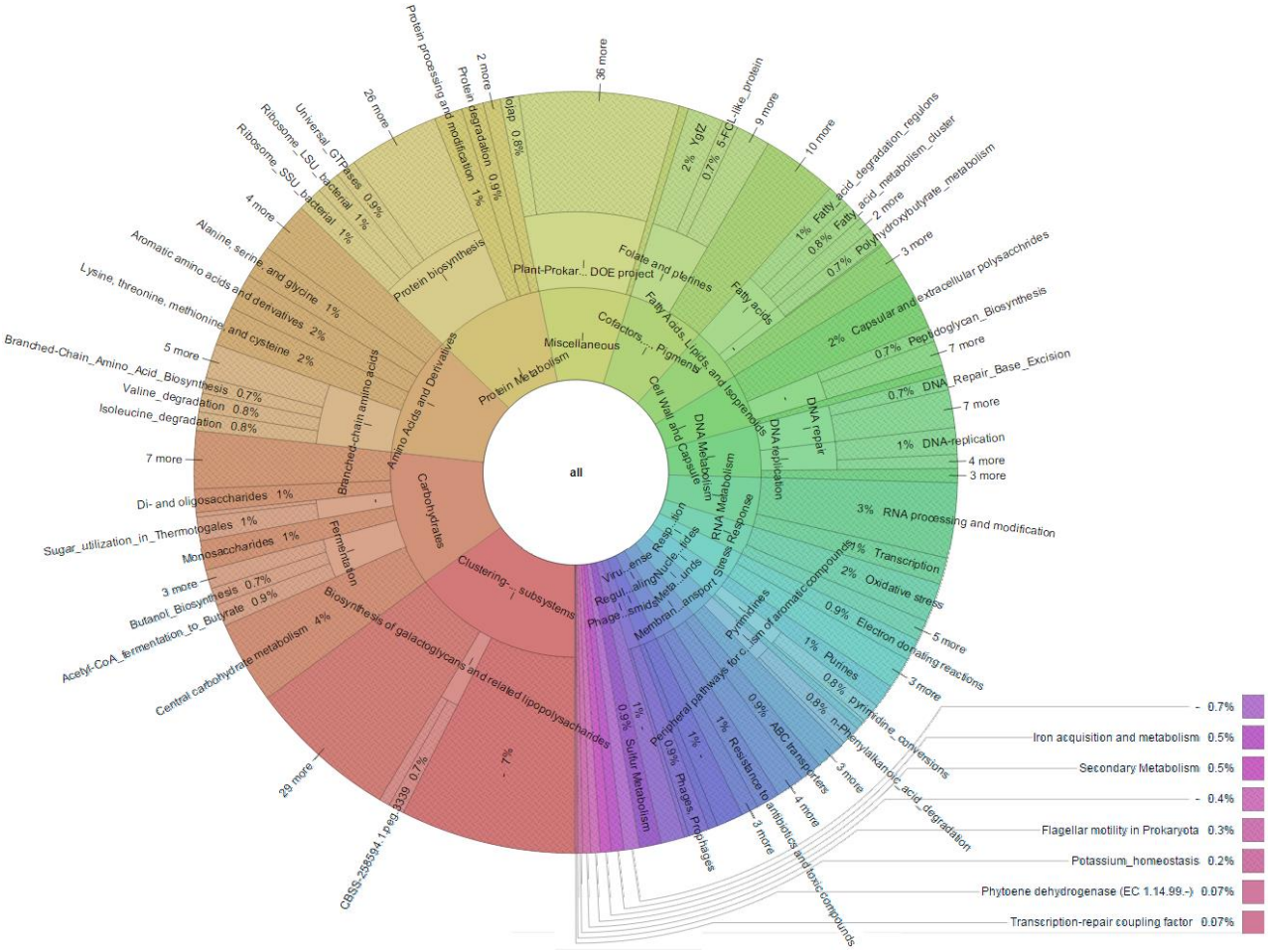


Figure 3.SI.8 MG RAST Krona high-level view of TM7-rich metagenome metabolic potential. Sequences were submitted to MGRAST with default sequence artefact removal, quality filtering and ambiguity filtering turned off (as already done).

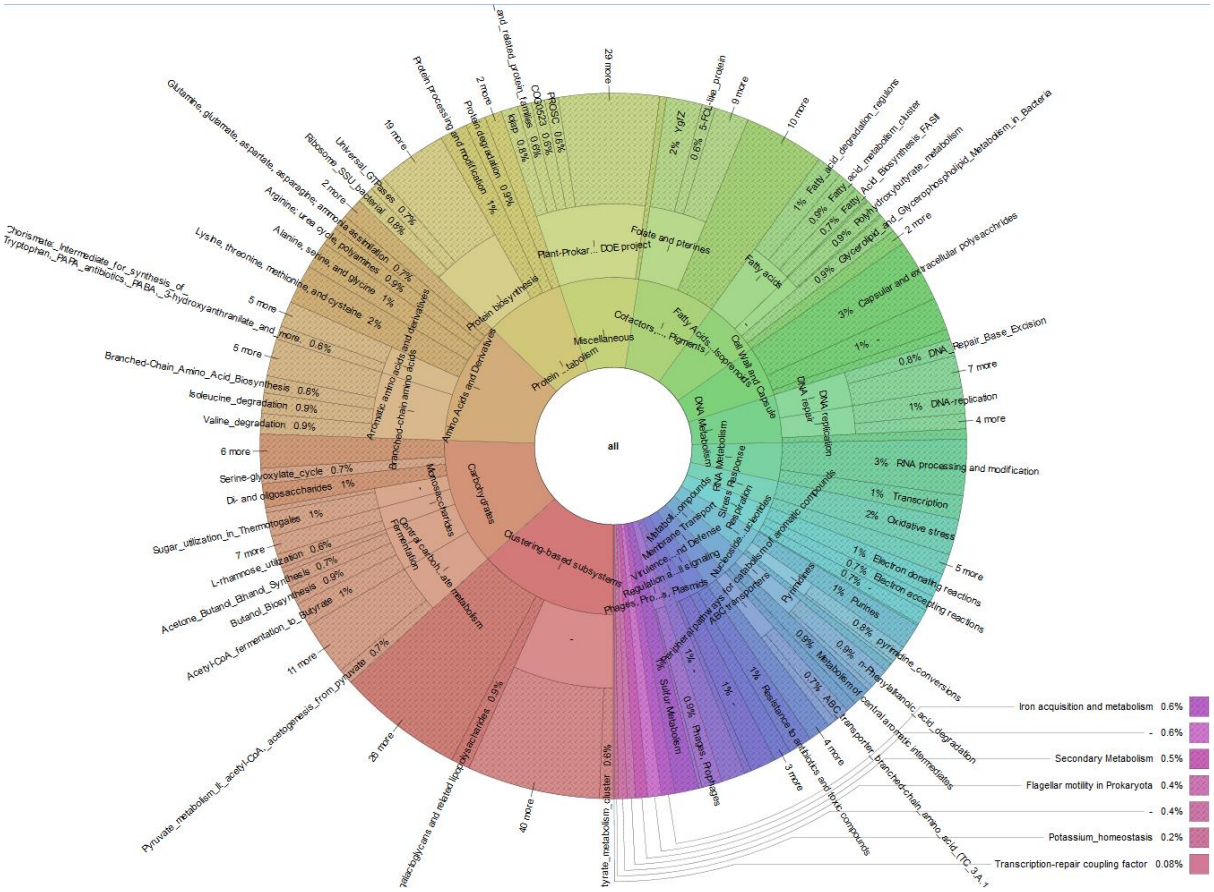


Figure 3.SI.9 MG RAST Krona high-level view of the metabolic potential of the TM7-rich metagenome after TM7 removal. Sequences were submitted to MGRAST with default sequence artefact removal, quality filtering and ambiguity filtering turned off (as already done).

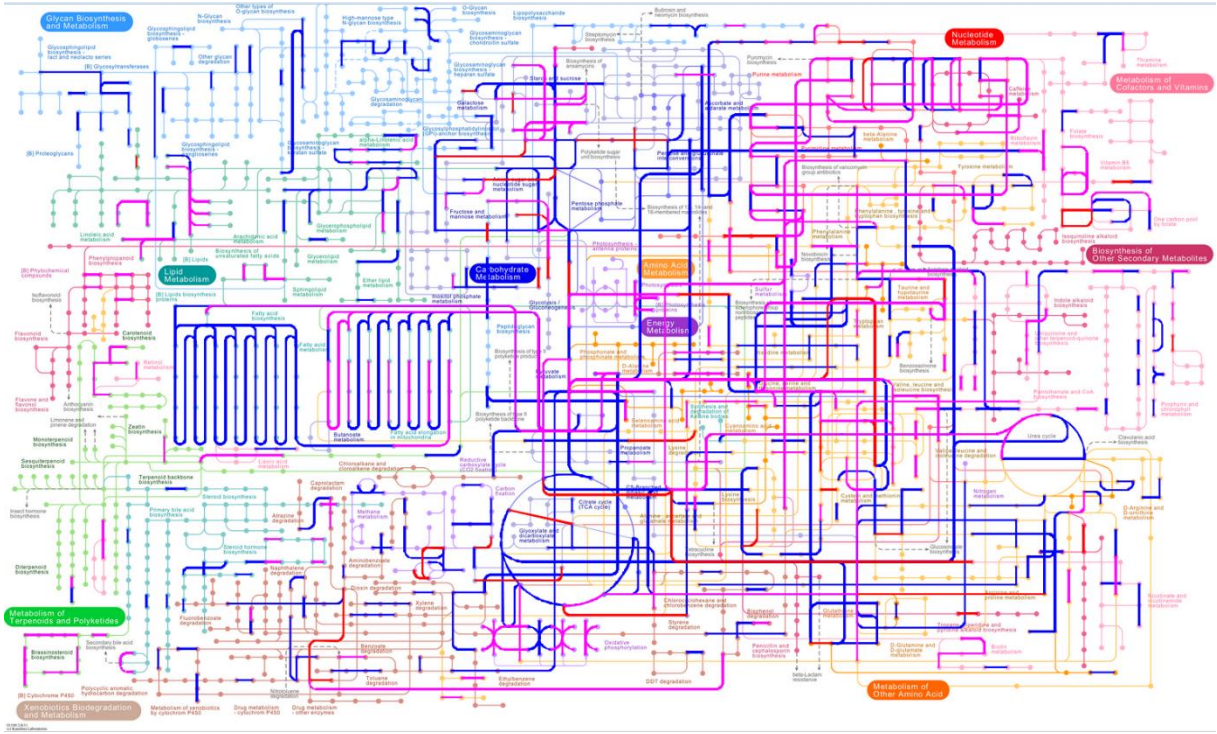


Figure 3.SI.10 Comparison between the TM7 containing metagenomic community predicted functions (red) and bl28ba (blue), with overlapping data (purple):

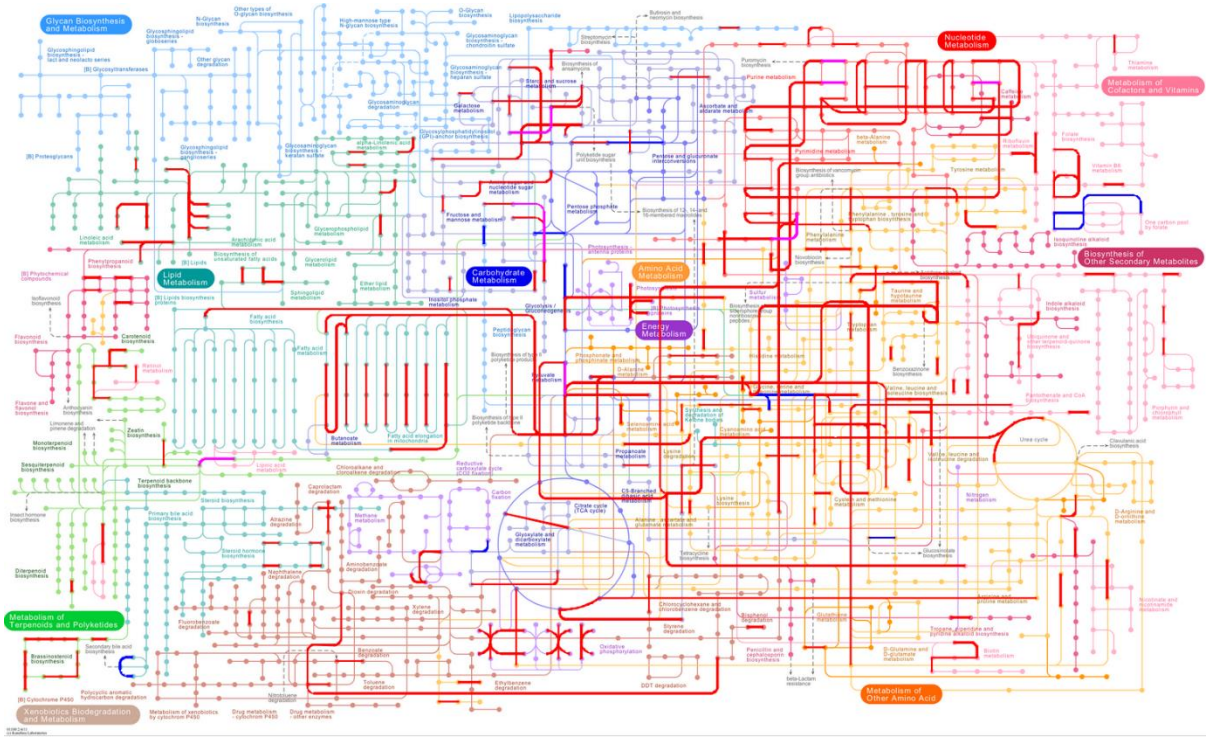


Figure 3.SI.11 Comparison between the TM7 (blue), and the non TM7 (red) sequence function predictions from the TM7 containing microcosm (purple = overlapping)
TM7 only adds functionality to the community in the one-carbon-pool, and activity with carbon-nitrogen bonds (other than peptide), possible role in carbon fixation

2.AX: Chapter 2 Appendix

2.AX.1 Development of a concentration device for volatile hydrocarbons.

In order to analyse low-concentration volatiles by Gas Chromatography with a Flame-Ionisation Detector, a concentrating device was required. The production of a device suitable for the requirements required several iterations of method development and testing.

For the trap specifications used in the 2015 in situ sampling, see Design 3.

Design 1.

The initial design consisted of a system where a sample stored in a previously evacuated serum vial with a silicon septum can be connected to a stream of nitrogen by a needle inserted through the septum. This causes flushing of the vial through a second needle through the septum, and into the first valve. In the “trap” position, this directs the flow of nitrogen containing the sample through a loop, held at -170°C with liquid nitrogen. At -170°C the volatiles, including isoprene, condense onto the loop, and are immobilised. The stream then flows through the second valve to the exhaust, where the flow rate can be monitored. In the trap position the third valve directs the input from the sampling loop of the gas chromatograph back into the gas chromatograph (Figure 2.AX.1).

Once the desired flow-through is achieved, to analyse the sample, the valves are switched to the release position (bottom to top, to prevent flow cut-off). In the release position, the liquid nitrogen is removed, and replaced with boiling water, causing volatilisation of the volatiles and flow through to the GC column in the reverse direction to trapping in the loop, increasing peak sharpness.

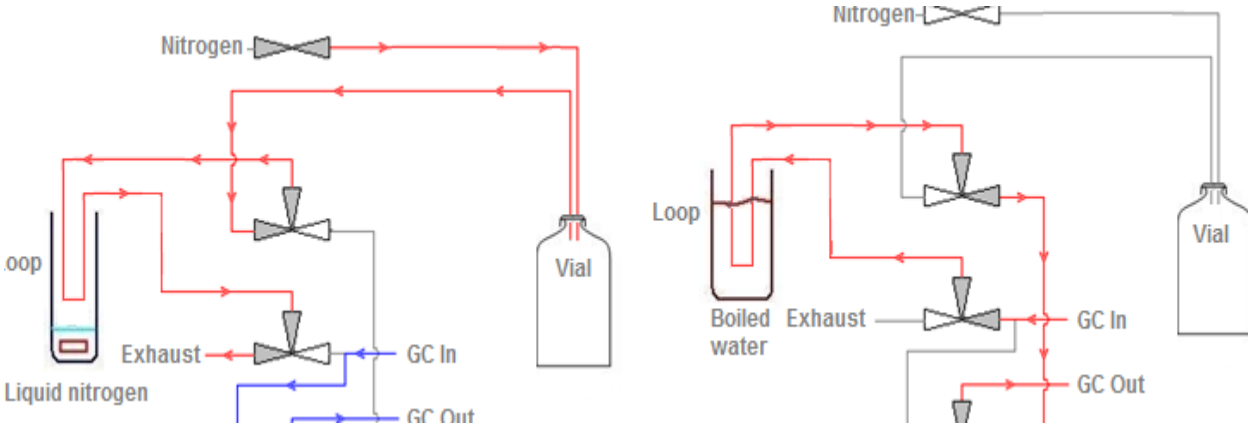


Figure 2.AX.1: Trap Design 1: left showing trap state, right showing release state.

The triple-valve system was chosen over the traditional 6-port valve system due to substantial cost savings.

Figure 2.AX.2 Photograph of Design 1 cryotrap.

Design 1 (Figure 2.AX.2) performed well on individual samples, however the in-line nature of the trap, without substantial moisture trapping capabilities caused a slow partial loop blockage on repeated trapping, causing pressure spikes, and

errors, in addition to which, the changing sampling design called for a more mobile solution.

Design 2.

The second phase of design consisted of a mobile-trapping system, where samples, typically of 20 cm^3, can be taken from the field, and passed through a loop cooled by liquid nitrogen, using negative pressure from suction generated from a second syringe, preventing leaks. Due to potential of nitrogen condensation, the loop was encased in an aluminium casing, allowing the high thermal conductivity of the aluminium to equilibrate the temperature between the temperature of the liquid nitrogen and the atmospheric temperature, with temperature adjustments made by altering the immersion depth, and temperature monitoring using a RS components APPA 5511 615-8212 Digital thermometer, with K-Type probes terminating above and below the loop internal to the casing and manufacturer calibrated down to -200°C.

Once the sample had passed through the loop, the “suction” syringe was re-routed to the output, and the line, as well as the output vessel (a small vial or syringe) was evacuated. Following this, the loop was sealed, and moved to hot water to re-volatise the volatiles, and then released into the evacuated output vial. The “suction” syringe could then be removed, allowing air follow-through to complete the trapping process (Figure 2.AX.3 for schematic, Figure 2.AX.4 for image).

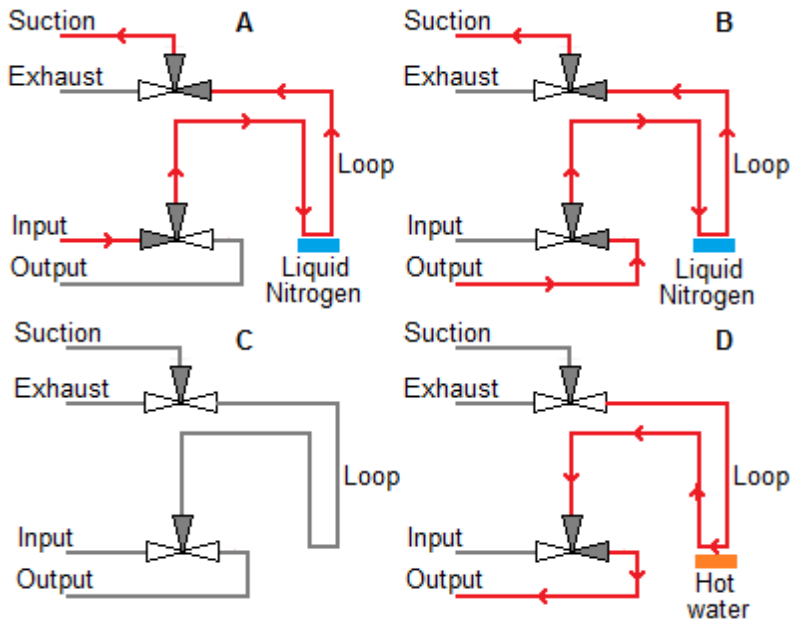


Figure 2.AX.3: Schematic of Trap design 2, showing A) Passing sample through loop using negative pressure, B) Evacuation of the Output line, C) removal from the liquid Nitrogen, D) Re-volatilisation and sample removal.

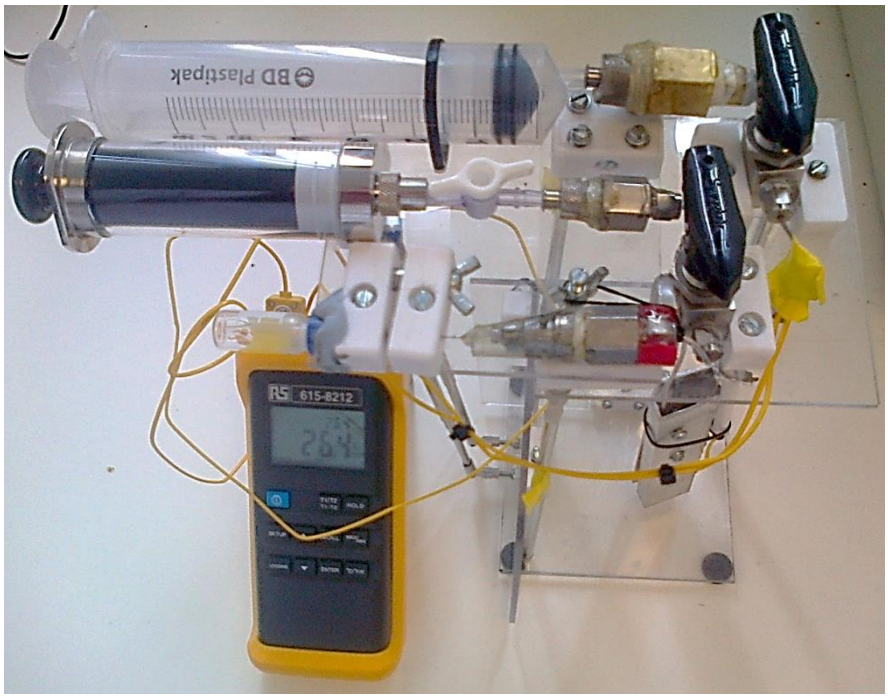


Figure 2.AX.4: Trap Design 2, showing trap and temperature probe.

Design 2 was able to trap up to 50 × initial concentration, based on 20 cm^3 gas samples, concentrated into 0.25 cm^3 with a flow rate of ~1 cm^3s^{-1}. However, trapping was not always entirely successful, necessitating a requirement for

analysis of the data in reference to an inert, volatile with similar properties. Despite this success, the changing requirements due to speed of degradation preventing sufficient sample numbers due to a long trapping time combined along with the relative complexity of operation, meant that a non-mobile based system again became preferable.

Design 3.

The cryotrap used for analysis of field experimental samples for concentration of gases comprised of a system in which the contents of a Tedlar bag could be moved through a loop, using negative pressure (suction). The loop was held at -160 using liquid nitrogen and a custom thermostatically controlled liquid nitrogen boiler (Built by UEA), causing condensation of volatile hydrocarbons. The Tedlars bags were kept on-ice during the evacuation stages, limiting water vapour. Once the volatiles were trapped in the loop, the line, and a small (0.250 cm^3) gas-tight syringe (attached to sample out, Figure 2.AX.5) were evacuated. Sealing of the line, and heating of the loop allowed re-volatilisation of the trapped gasses, and opening the line to the sample out allowed the hydrocarbons to enter the syringe. Opening of the ‘Suction’ port (Figure 2.AX.5) resulted in air being pulled through into the syringe, carrying more volatiles with it. A small movement in of the ‘suction’ syringe at this point would confirm successful evacuation of volatiles. The syringe, in this case, a SGE Gastight locking leur-lock syringe, could be used for directly injecting into a GC-FID.

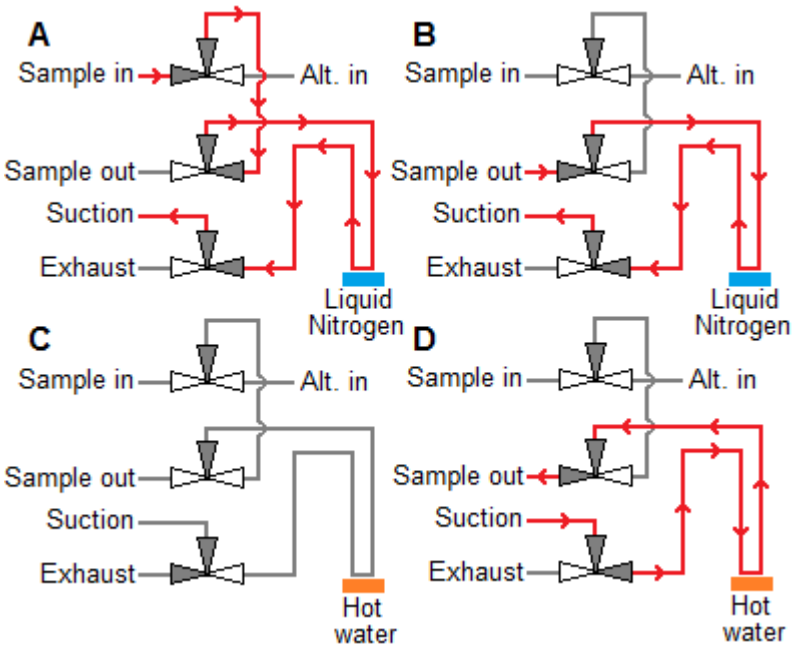


Figure 2.AX.5: Schematic of Trap design 2, showing A) Volatile trapping, B) Application of vacuum to line, C) Volatisation of volatiles, D) Removal of volatiles.

The piping consisted of 1/16” SS316 Stainless steel Swagelok tubing, with a 0.02 mm wall, giving a low reactivity, cost and internal diameter. Although a low volume is advantageous for evacuation, in order to increase trapping efficiency, and condensation site specific volume, the loop was tightly wound with around 0.5 m of tubing incorporated.

Attachment to the 3/16 Tedlar bag outlet was completed by a 1/16” to 3/16” union, with the 3/16” union tightened around the ferrule to give a close seal to the interchanging 3/16” bag outlets. Steel 1/16” to female-leur-lock adapters were purchased from ColeParmer for the suction terminal, and VICI Valco for the sample-out terminal (due to low dead-volume). All other unions and joints were purchased in SS316 stainless steel from Swagelok. Valves were Swagelok Stainless steel 1-piece ball valves from the 40G s3 series.

The case was constructed using clear polycarbonate sheeting, with easy access panels to the top 2/3 of the back, and the bottom third of the front, for easy visibility, construction and leak testing. The case was designed to be as small as possible, whilst being stable and durable (Figure 2.AX.6). The case also contained adaptors for headspace flushing, if required, through the Alt-in (Figure 2.AX.5), and an additional Nitrogen valve.

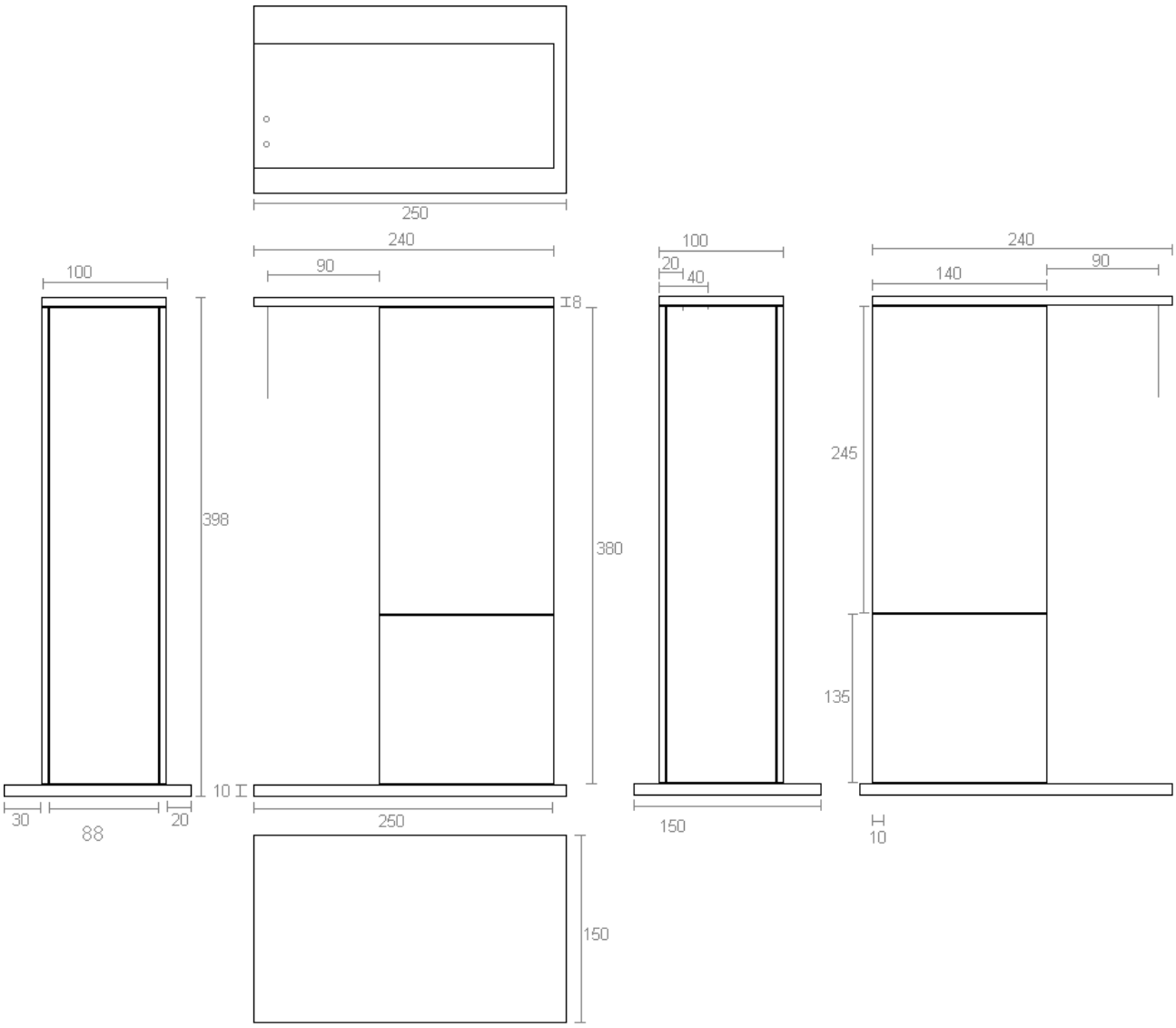


Figure 2.AX.6: Cryotrap case dimensions (mm)

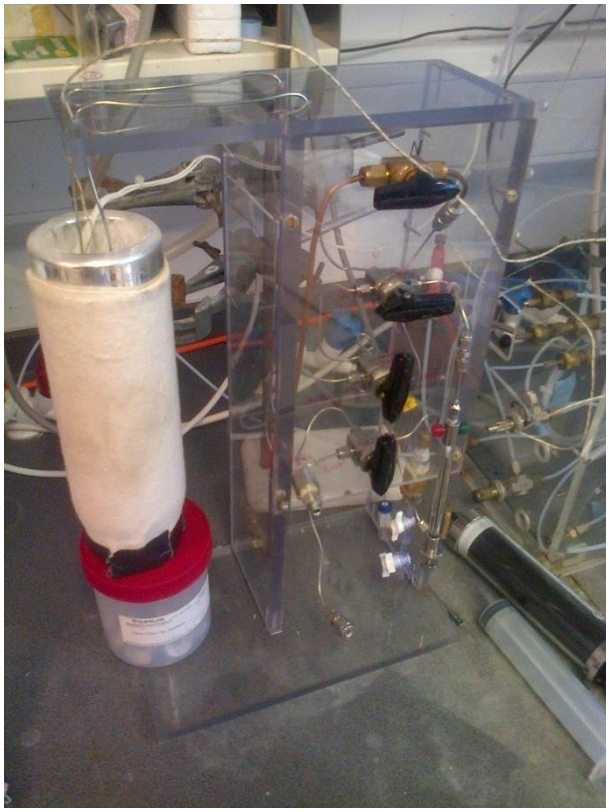


Figure 2.AX.7: Cryotrap preparing for stage D, hot water volatilisation (Figure 2.AX.5).

The trap was showing to be effective at concentrating hydrocarbons, with often over 200 × initial concentration of isoprene (based on 50 cm^3 samples at 60 cm^3 min^{-1}, maximum theoretical efficiency 400 ×), with a linear relationship between Isoprene and the Decaflouropentane standards on different trapping efficiencies, demonstrating that the occasional poorly-trapped sample does not affect ratio-based analysis.

2.AX.2 Images of soil volatile chambers

Figure 2.AX.8 Image of soil chambers for measuring changes in isoprene flux. Left: Chambers, Middle: Chambers deployed (open), Right: Schematic of soil chamber.

3.AX: Chapter 3 Appendix

Table 3.AX.1 description of isoprene degrading strains.

Name	Closest	*isoA*	*16S rRNA*	Description	Degradation	Growth	Soil /Leaf
W1ab	*R. wratislaviensis*	✓	✓	1.5 mm, Circular, Cream, Raised, Smooth, Mucoid, Filiform.			L
104bca	*R. globerulus*	✓	✓	3.5 mm, Circular, Transparent, Umbonate, Shiny, Moist, Lobate.			S
bW1	*R. globerulus*	✓	✓	3 mm, Circular, Cloudy Dark Orange, Convex, Smooth, Moist, Entire.			S
TC25ba	*P. panaciterrae*	✓	✓	4 mm, Irregular, Cloudy Pale White, Convex, Glistening, Moist, Entire.			L
W1a	*R. wratislaviensis*	✓	✓	3 mm, Circular, Opaque Orange, Raised, Smooth, Moist, Entire. Insperable from a fungi			L
2(11b)	*Rhodococcus erythropolis*	✓	✓	2.5 mm, Circular, Transparent, Flat, Veined, Dry, Undulate.			S
U8b	*R. erythropolis*	✓	✓	7 mm, Circular, Opaque Pale Yellow, Raised, Shiny, Dry, Lobate.			S
Tc27b	*P. panaciterrae*	✓	✓	6 mm, Irregular, Cloudy Cream, Convex, Smooth, Moist, Entire.			S
52q	*Devosia riboflavin*	✓	✓	4 mm, Irregular, Opaque White, Flat, Dull, Moist, Entire.			L
W9	*R. fascians*	✓	✓	1 mm, Circular, Opaque Orange, Raised, Smooth, Butyrous, Entire.	✓	✓	L
P10dc	*R. wratislaviensis*	✓	✓	1 mm, Circular, Cream, Raised, Smooth, Moist, Entire.			S
6(1b)	*Aeromicrobium ginsengisoli*	✓	✓	4 mm, Irregular, Pale Cream, Umbonate, Dull, Moist, Lobate.	✓	✓	L
O6	*R. wratislaviensis*	✓	✓	1 mm, Circular, Opaque Yellow, Raised, Smooth, Butyrous, Entire.			L
GW7	*Pseudomonas mohnii*	✓	✓	1 mm, Circular, Transparent, Raised, Shiny, Moist, Entire.			L
P1	*Pseudomonas fragi*	✓	✓	1 mm, Circular, Opaque Pale Yellow, Umbonate, Smooth, Dry, Lobate.			L
GM3	*Methylobacterium mesophilicum*	✓	✓	1 mm, Circular, Pale Pink, Raised, Shiny, Moist, Entire.	✓	✓	S
5(2a)b	*R. erythropolis*	✓	✓	3 mm, Irregular, Transparent, Flat, Dull, Moist, Lobate.			S
32d-2	*R. erythropolis*	✓	✓	3 mm, Circular, Opaque Pale Cream, Raised, Shiny, Butyrous, Curled. Kanamycin resistant.			S
W2	*Methylobacterium bullatum*	✓	✓	1 mm, Circular, Pale Pink, Raised, Smooth, Moist, Entire.			L
12.88	*Nocardia takedensis*	✓	✓	1 mm, Circular, Opaque White, Raised, Rough, Dry, Lobate.			S
P12	*A. nitroguajacolicus*	✓	✓	6 mm, Circular, Pale Yellow, Crateriform, Shiny, Moist, Curled.	✓		L
bA2c	*R. globerulus*	✓	✓	3 mm, Circular, Pale Orange, Flat, Dull, Dry, Curled.			S

Name	Closest	*isoA*	*16S rRNA*	Description	Degradation	Growth	Soil /Leaf
8c1c	*Rhodococcus globerulus*	✓	✓	2 mm, Circular, Opaque Pale Cream, Raised, Rough, Dry, Curled.			S
b22P25		✓		1 mm, Circular, Translucent Pale White, Flat, Rough, Moist, Lobate.	✓	✓	S
bE1	*R. globerulus*	✓	✓	4 mm, Circular, Cloudy Dark Orange, Raised, Smooth, Moist, Lobate.	✓		S
32d-3c	*R. erythropolis*	✓	✓	2.5 mm, Circular, Translucent White, Umbonate, Dull, Dry, Lobate.	✓	✓	S
1(2b)b	*R. erythropolis*	✓	✓	7 mm, Circular, Transparent, Flat, Veined, Moist, Filiform. Kanamycin resistant.			S
W7	*S. rhizophila*	✓	✓	1 mm, Circular, Yellow, Raised, Shiny, Moist, Entire.			L
U7b	*R. erythropolis*	✓	✓	1 mm, Circular, Opaque Cream, Raised, Shiny, Moist, Entire.	✓	✓	S
2(11a)	*Rhodococcus sp.*	✓	✓	2 mm, Circular, Pale White, Raised, Smooth, Moist, Entire.			S
5(2a)e	*R. erythropolis*	✓	✓	3.5 mm, Rhizoid, Cloudy Dark White, Raised, Rough, Mucoid, Entire.			S
4(2c)	*R. erythropolis*	✓	✓	3 mm, Circular, Cloudy Pale Orange, Convex, Shiny, Butyrous, Curled. Kanamycin resistant			S
TC4e2	*R. wratislaviensis*	✓	✓				S
Tc4jb2	*R. wratislaviensis*	✓	✓	2 mm, Circular, Opaque Cream, Convex, Smooth, Moist, Entire.			S
b22c	*R. erythropolis*	✓	✓	11 mm, Irregular, Translucent Pale White, Flat, Veined, Viscous, Entire.			S
B1	*R. erythropolis*	✓	✓	10 mm, Circular, Pale White, Flat, Rough, Dry, Entire.	✓		L
32d-1	*R. erythropolis*	✓	✓				S
P10a	*R. cercidiphylli*	✓	✓	3 mm, Circular, Translucent Pale Yellow, Raised, Smooth, Moist, Entire.			L
b22T6G24b	*R. erythropolis*	✓	✓	14 mm, Irregular, Opaque Orange, Convex, Wrinkled, Viscous, Lobate.			S
bL28a	*R. erythropolis*	✓	✓	3.5 mm, Circular, Opaque Pale Orange, Raised, Smooth, Viscous, Curled.	✓		S
1(2b)bb	*R. erythropolis*	✓	✓	1.5 mm, Circular, Translucent Pale White, Flat, Dull, Dry, Filiform.			S
bl28b	*R. erythropolis*	✓	✓	2.5 mm, Circular, Translucent Light Grey, Umbonate, Dull, Moist, Curled.	✓	✓	S
32d-3b	*R. erythropolis*	✓	✓	3.5 mm, Irregular, Cloudy White, Flat, Dull, Dry, Filiform.			S
P8	*R. fascians*	✓	✓	2 mm, Circular, Pale Yellow, Convex, Shiny, Moist, Entire.			L
U6ba	*Mezorhizobium*	✓	✓	2 mm, Circular, Pale Orange, Raised, Dull, Dry, Entire.			S
TC4cb2	*R. wratislaviensis*	✓	✓				S
U8	*R. erythropolis*	✓	✓	1.2 mm, Circular, Opaque Orange, Pulvinate, Shiny, Moist, Entire.			S
W1aba	*R. wratislaviensis*	✓	✓	3 mm, Rhizoid, Opaque White, Flat, Rough, Dry, Filiform.			L

Name	Closest	*isoA*	*16S rRNA*	Description	Degradation	Growth	Soil /Leaf
b228b	*R. erythropolis*	✓	✓	6 mm, Irregular, Iridescent White, Flat, Wrinkled, Moist, Entire.			S
bE1b	*Rhodococcus sp.*	✓	✓	3.5 mm, Circular, Translucent Pale White, Flat, Dull, Moist, Entire.	✓		S
A,AgA	*Variovorax boronicumulans*	✓	✓	Punctiform mm, Circular, Transparent, Umbonate, Dull, Dry, Lobate.			S
Tc25b	*Pedobacter panaciterrae*	✓	✓	3 mm, Circular, Cloudy Dark Yellow, Convex, Glistening, Moist, Entire.			S
B22ch3	*Ensifer adhaerens*	✓	✓	2.5 mm, Circular, Transparent, Convex, Glistening, Moist, Entire.			S
P7	*A. nitroguajacolicus*	✓	✓	2 mm, Circular, Transparent, Raised, Shiny, Moist, Entire.			L
T82EG2	*R. globerulus*	✓	✓	1 mm, Circular, Cloudy Pale White, Raised, Smooth, Moist, Entire.			S
BE156b	*Rhodococcus sp.*	✓	✓	1.5 mm, Circular, White, Umbonate, Rough, Moist, Lobate.			S
3.88	*R. globerulus*	✓	✓	2 mm, Circular, Opaque Pale White, Raised, Rough, Dry, Filiform.			S
S2s	*Rhodococcus sp.*	✓	✓	4 mm, Circular, Transparent, Umbonate, Veined, Moist, Filiform. Grows as well without isoprene.	✗		S
bA1A	*Pimelobacter simplex*	✓	✓	0.5 mm, Circular, White, Raised, Smooth, Moist, Entire.		✓	S
5b2a	*Rhodococcus sp.*	✓	✓	3 mm, Irregular, Cloudy Pale Orange, Convex, Shiny, Butyrous, Curled. Kanamycin resistant.			S
bl3bb	*R. erythropolis*	✓	✓	1.5 mm, Circular, Opaque White, Flat, Smooth, Moist, Entire.	✓	✓	S
R2A104a	*Rhodococcus sp.*	✓	✓	6 mm, Circular, Translucent Pale White, Raised, Shiny, Moist, Lobate.			S
bl3a	*R. erythropolis*	✓	✓	2 mm, Circular, Translucent Dark Cream, Convex, Shiny, Mucoid, Lobate.	✓	✓	S
bl3b	*R. erythropolis*	✓	✓	4 mm, Circular, Cloudy Pale Cream, Raised, Shiny, Moist, Entire.			S
5c(2a)	*R. erythropolis*	✓	✓	3 mm, Circular, Opaque Light Orange, Umbonate, Smooth, Moist, Lobate. Kanamycin resistant.			S
T3EG15a	*R. erythropolis*	✓	✓	4 mm, Circular, Opaque Pale White, Flat, Smooth, Dry, Lobate.			S
5b(2a)	*R. erythropolis*	✓	✓	3 mm, Irregular, Cloudy Cream, Convex, Shiny, Moist, Curled.			S
G28	*Nitrobacter vulgaris*	✓	✓	Punctiform mm, Circular, Transparent, Pulvinate, Shiny, Moist, Entire.			S
P13	*Arthrobacter nitroguajacolicus*	✓	✓	1 mm, Circular, Translucent Pale Yellow, Raised, Smooth, Moist, Entire.			L
104bc	*R. globerulus*	✓	✓	4 mm, Circular, Transparent, Umbonate, Shiny, Moist, Lobate.			S
bl28ba	*R. erythropolis*	✓	✓	5 mm, Circular, Pale Cream, Raised, Shiny, Moist, Lobate.	✓	✓	S
C:IT		✓	✓	1.5 mm, Irregular, Translucent Light White, Raised, Glistening, Moist, Entire.	✓	✓	S
GM7	*R. fascians*	✓	✓	8 mm, Circular, Opaque Yellow, Umbonate, Shiny, Moist, Entire.			L
104bcab	*R. globerulus*	✓	✓	6 mm, Circular, Translucent Pale White, Raised, Shiny, Moist, Lobate.			S

Name	Closest	*isoA*	*16S rRNA*	Description	Degradation	Growth	Soil /Leaf
1(2b)b2	*R. erythropolis*	✓					S
P3	*A. nitroguajacolicus*	✓	✓	2.5 mm, Circular, Translucent Dark Orange, Raised, Smooth, Moist, Lobate.			L
TC13b	*Bosea robiniae*	✓	✓	3.5 mm, Circular, Dark Cream, Umbonate, Smooth, Moist, Curled.	✓	✓	S
P1a	*Rhodococcus fascians*	✓	✓	2 mm, Circular, Opaque Orange, Raised, Smooth, Butyrous, Entire.			L
GM3b	*Rhodococcus sp.*	✓	✓	Punctiform mm, Circular, Translucent Pale Pink, Raised, Shiny, Moist, Entire.			L
b222P1	*R. globerulus*	✓	✓	4 mm, Circular, White, Raised, Wrinkled, Moist, Entire.			S
bl3ba	*Stenotrophomonas rhizophila*	✓	✓	2.5 mm, Circular, Cloudy Dark White, Convex, Glistening, Moist, Entire.	✓	✓	S
bW2b	*Rhodococcus sp.*	✓	✓	2 mm, Irregular, Dark Orange, Umbonate, Smooth, Moist, Entire.			S
P10c-A	*R. wratislaviensis*	✓	✓	4 mm, Irregular, Opaque Orange, Umbonate, Dull, Moist, Filiform.			L
P102-Ab	*R. wratislaviensis*	✓	✓	1.5 mm, Circular, White, Raised, Smooth, Viscous, Entire.			L
P6	*Rhodococcus sp.*	✓	✓	0.5 mm, Circular, Transparent, Raised, Shiny, Moist, Entire.			L
7.88				4 mm, Irregular, Opaque Pale Orange, Umbonate, Smooth, Moist, Curled.			S
104102A				2 mm, Irregular, Cloudy White, Convex, Shiny, Moist, Entire.	✓	✓	S
104bcaa	*R. globerulus*						S
104bcac	*R. globerulus*						S
104bcd				3 mm, Circular, Translucent Pale Cream, Crateriform, Shiny, Moist, Lobate.			S
104c				3 mm, Circular, Cloudy Pale Orange, Convex, Smooth, Moist, Filiform.			S
104ca				3 mm, Circular, Cloudy Pale Orange, Convex, Smooth, Moist, Filiform.			S
104cc				4 mm, Circular, Translucent Pale White, Umbonate, Rough, Dry, Lobate.	✓	✓	S
104cd				4 mm, Circular, Translucent White, Flat, Rough, Dry, Lobate.	✓	✓	S
104No2				1 mm, Circular, Translucent Pale White, Flat, Rough, Moist, Lobate.	✓	✓	S
211aa				20 mm, Irregular, Transparent, Flat, Dull, Moist, Entire.	✓	✓	S
211ab				3 mm, Circular, Opaque Cream, Raised, Dull, Dry, Filiform.			S
24C1a				3.5 mm, Circular, Translucent Pale White, Umbonate, Wrinkled, Dry, Entire. kanamycin resistant			S
24C2a				3 mm, Irregular, Transparent, Flat, Dull, Moist, Lobate.			S
25bU3c				1 mm, Circular, Translucent Pale Yellow, Raised, Glistening, Moist, Entire.			S

Name	Closest	*16S rRNA isoA*	Description	Degradation	Growth	Soil /Leaf
32d	*R. erythropolis*					S
32d-1b			2 mm, Circular, Opaque Orange, Raised, Smooth, Moist, Lobate.			S
32d-3	*R. erythropolis*		3 mm, Irregular, Translucent Pale Grey, Flat, Rough, Dry, Filiform.			S
52ab			3 mm, Circular, Opaque Pale Cream, Umbonate, Shiny, Moist, Curled.	✓	✓	S
5b(2a)2	*R. erythropolis*					S
8C1a			2 mm, Circular, Opaque Light Orange, Convex, Shiny, Mucoid, Entire.			S
8c1b			2 mm, Circular, Opaque Pale Orange, Convex, Smooth, Moist, Entire.			S
8C2a			2 mm, Circular, Translucent Pale Yellow, Umbonate, Veined, Moist, Lobate.			S
B2_28ba			0.5 mm, Circular, Pale Orange, Pulvinate, Dull, Dry, Entire.	✓	✓	L
b2_3a			2 mm, Circular, White, Convex, Smooth, Moist, Entire.			S
b22b			1.5 mm, Filamentous, Translucent Pale White, Pulvinate, Rough, Dry, Entire.	✓	✓	S
b22m			8 mm, Irregular, Cloudy Dark White, Crateriform, Wrinkled, Viscous, Curled.	✓	✓	S
b22p116			3 mm, Circular, Translucent Pale Cream, Umbonate, Smooth, Moist, Undulate.	✓	✓	S
b22P16			2 mm, Circular, Opaque White, Convex, Smooth, Moist, Entire.	✓	✓	S
b22p5			4 mm, Circular, Translucent White, Umbonate, Smooth, Dry, Lobate.	✓	✓	S
b22T3G9b			1 mm, Circular, Translucent Pale White, Flat, Rough, Dry, Filiform.	✓	✓	S
b22T4EG19			1.5 mm, Circular, Opaque Pale Orange, Raised, Smooth, Moist, Entire.			S
b22T8G20			2 mm, Circular, Translucent Pale White, Raised, Smooth, Moist, Undulate.			S
b22T8G2Ob			2 mm, Circular, Opaque Light Orange, Convex, Smooth, Moist, Entire.			S
b23b			5 mm, Irregular, Transparent, Flat, Dull, Dry, Entire.			S
b310			3 mm, Irregular, Cloudy Pale Cream, Crateriform, Shiny, Butyrous, Entire.			S
bd28b			4 mm, Circular, Cloudy Light Cream, Umbonate, Rough, Dry, Filiform.			S
bE156			18 mm, Circular, Cloudy Pale White, Flat, Rough, Dry, Entire.			S
bl3bc			3 mm, Circular, Cloudy Cream, Raised, Glistening, Mucoid, Entire.	✓	✓	S
bl3bd			4 mm, Circular, Translucent White, Umbonate, Rough, Dry, Lobate.	✓	✓	S
bp110			2 mm, Circular, Translucent Pale White, Umbonate, Rough, Moist, Lobate.	✓	✓	S
bp316			3.5 mm, Circular, Translucent White, Raised, Shiny, Moist, Entire.	✓	✓	S

Name	Closest *16S rRNA isoA*	Description	Degradation	Growth	Soil /Leaf
BW2		3 mm, Circular, Cloudy Pale White, Raised, Shiny, Moist, Entire.			S
DoT6		5 mm, Circular, Opaque Pale Yellow, Umbonate, Rough, Viscous, Entire.			L
DoT7		10 mm, Circular, Opaque Pale Cream, Convex, Rough, Viscous, Lobate.			S
G28b		1 mm, Circular, Transparent, Flat, Dull, Dry, Lobate.	✓	✓	S
GM6		0 mm, Circular, Translucent Pale Orange, Raised, Shiny, Moist, Entire.			L
GM8		Punctiform mm, Circular, Pale Yellow, Raised, Shiny, Moist, Entire.			L
L1		1 mm, Circular, Pale Cream, Raised, Smooth, Moist, Entire.			P
L2		0.5 mm, Circular, Pale Yellow, Raised, Glistening, Moist, Entire.			P
LI9b		20 mm, Irregular, Transparent, Flat, Dull, Moist, Entire.	✓	✓	S
LI9		20 mm, Irregular, Transparent, Flat, Dull, Moist, Entire.	✓	✓	S
O2		1 mm, Circular, Opaque Pale Orange, Raised, Smooth, Moist, Entire.	✓		L
O4		0.5 mm, Circular, Pale Orange, Raised, Smooth, Moist, Entire.	✓		L
P102 Aa		2.5 mm, Rhizoid, Opaque Light Cream, Raised, Rough, Dry, Entire.	✓	✓	L
P10b		1.5 mm, Circular, Cream, Raised, Smooth, Moist, Entire.			L
P10c b		2 mm, Circular, Opaque Pale Orange, Convex, Smooth, Moist, Entire.			L
P10c -ba		4 mm, Rhizoid, Opaque Pale Cream, Raised, Dull, Dry, Entire.	✓	✓	L
P10d -c		0.4 mm, Circular, Pale Yellow, Raised, Shiny, Moist, Entire.			L
P2		0.5 mm, Circular, Yellow, Raised, Smooth, Moist, Entire.	✓		L
P2-b		6 mm, Irregular, Transparent, Flat, Dull, Dry, Entire.	✓	✓	L
T3E G15b		2 mm, Filamentous, Transparent, Raised, Shiny, Brittle, Entire.			S
T42C G21a		2 mm, Circular, Pale White, Convex, Smooth, Moist, Entire.			S
T42E G21b		1.5 mm, Circular, Transparent, Flat, Dull, Dry, Undulate.	✓	✓	S
T42E G26b		3 mm, Circular, Transparent, Flat, Dull, Dry, Lobate.	✓	✓	S
TC27 cd		0.5 mm, Circular, Translucent Pale White, Flat, Rough, Moist, Filiform.	✓	✓	S
TC4		3 mm, Circular, Transparent, Raised, Shiny, Mucoid, Entire.			S
Tc4c 2		3 mm, Circular, Cloudy Cream, Convex, Shiny, Moist, Entire.			S
TC4f b2		1.5 mm, Circular, Translucent Pale Cream, Raised, Shiny, Moist, Entire.			S

Name	Closest	*16S rRNA*	*isoA*	Description	Degradation	Growth	Soil /Leaf
U31				1 mm, Circular, Cloudy Cream, Convex, Shiny, Moist, Entire.			S
U3b				8 mm, Circular, Translucent Pale White, Flat, Rough, Moist, Lobate.	✓	✓	S
U6bb				1 mm, Circular, Transparent, Flat, Smooth, Moist, Entire.	✓	✓	S
U7a				3.5 mm, Circular, Pale White, Raised, Smooth, Moist, Curled.			S
W1a	*R. wratislaviensis*			1 mm, Circular, Opaque Cream, Convex, Smooth, Moist, Filiform.			L
ba2							
W5				2.5 mm, Circular, White, Raised, Smooth, Moist, Undulate.	✓	✓	L
W6				3 mm, Circular, Cloudy Cream, Raised, Shiny, Moist, Entire.			L

Note: list is incomplete as not all strains were described. S = Soil, L = Leaf, P = lake water. Tick marks in the growth and degradation columns are present where tested and positive, crosses are present where tested and negative, no marks indicate that the strain was not tested in that manner. Tick marks in the 16S rRNA and IsoA columns denote that the strain was sequenced for the gene and used in the core analysis which forms part of this chapter. Closest relative information was added where 16s rRNA sequences were obtained and over 97% similar in blast results. Degradation was tested in liquid culture, demonstrating a decrease in the isoprene concentration, growth was tested by comparing colony presence/size on minimal media plates in and out of isoprene. All isolates were cultured for at least seven rounds on minimal media washed agar plates with isoprene as the sole added carbon source.

3.AX.2 List of carbon sources representative strains were tested on:

Compounds representative strains were screened for utilisation of consisted of:

Custom plates with; Isoprene, Toluene, DMSO, Hexane, Ethane, Benzene, Propane, and Methane.

Biolog plates with: L-Arabinose, N-Acetyl-DGlucosamine, D-Saccharic Acid, Succinic Acid, D-Galactose, L-Aspartic Acid, L-Proline, D-Alanine, D-Trehalose, D-Mannose, Dulcitol, D-Serine, D-Sorbitol, Glycerol, L-Fucose, D-Glucuronic Acid, D-Gluconic Acid, D,L-a-GlycerolPhosphate, D-Xylose, L-Lactic Acid, Formic Acid, D-Mannitol, L-Glutamic Acid, D-Glucose-6- Phosphate, D-Galactonic Acid-g-Lactone, DL-Malic Acid, D-Ribose, Tween 20, L-Rhamnose, D-Fructose, Acetic Acid, a-D-Glucose, Maltose, D-Melibiose, Thymidine , L-Asparagine, D-Aspartic Acid, D-Glucosaminic Acid, 1,2-Propanediol, Tween 40,

a-Keto-Glutaric Acid, a-Keto-Butyric Acid, a-Methyl-DGalactoside, a-D-Lactose, Lactulose, Sucrose, Uridine, L-Glutamine, m-Tartaric Acid, D-Glucose-1-Phosphate, D-Fructose-6- Phosphate, Tween 80, a-Hydroxy Glutaric Acid-g-Lactone, a-Hydroxy Butyric Acid, b-Methyl-DGlucoside, Adonitol, Maltotriose, 2-Deoxy Adenosine, Adenosine, Glycyl-L-Aspartic Acid, Citric Acid, m-Inositol, D-Threonine, Fumaric Acid, Bromo Succinic Acid, Propionic Acid, Mucic Acid, Glycolic Acid, Glyoxylic Acid, D-Cellobiose, Inosine, Glycyl-LGlutamic Acid, Tricarballylic Acid, L-Serine, L-Threonine, L-Alanine, L-Alanyl-Glycine, Acetoacetic Acid, N-Acetyl-b-DMannosamine, Mono Methyl Succinate, Methyl Pyruvate, D-Malic Acid, L-Malic Acid, Glycyl-L-Proline, p-Hydroxy Phenyl Acetic Acid, m-Hydroxy Phenyl Acetic Acid, Tyramine, D-Psicose, L-Lyxose, Glucuronamide, Pyruvic Acid, L-Galactonic Acid-g-Lactone, D-Galacturonic Acid, Phenylethylamine, 2-Aminoethanol, Chondroitin Sulfate C, a-Cyclodextrin, b-Cyclodextrin, g-Cyclodextrin, Dextrin, Gelatin, Glycogen, Inulin, Laminarin, Mannan, Pectin, N-Acetyl-DGalactosamine, N-AcetylNeuraminic Acid, b-D-Allose, Amygdalin, D-Arabinose, D-Arabitol, L-Arabitol, Arbutin, 2-Deoxy-DRibose, i-Erythritol, D-Fucose, 3-0-b-D-Galactopyranosyl-DArabinose, Gentiobiose, L-Glucose, Lactitol, D-Melezitose, Maltitol, a-Methyl-DGlucoside, b-Methyl-DGalactoside, 3-Methyl Glucose, b-Methyl-DGlucuronic Acid, a-Methyl-DMannoside, b-Methyl-DXyloside, Palatinose, D-Raffinose, Salicin , Sedoheptulosan, L-Sorbose, Stachyose, D-Tagatose, Turanose, Xylitol, N-Acetyl-DGlucosaminitol, g-Amino Butyric Acid, d-Amino Valeric Acid, Butyric Acid, Capric Acid, Caproic Acid, Citraconic Acid, Citramalic Acid, D-Glucosamine, 2-Hydroxy Benzoic Acid, 4-Hydroxy Benzoic Acid, b-Hydroxy Butyric Acid, g-Hydroxy Butyric Acid, a-Keto-Valeric Acid, Itaconic Acid, 5-Keto-

DGluconic Acid, D-Lactic Acid Methyl Ester, Malonic Acid, Melibionic Acid, Oxalic Acid, Oxalomalic Acid, Quinic Acid, D-Ribono-1,4- Lactone, Sebacic Acid, Sorbic Acid, Succinamic Acid, D-Tartaric Acid, L-Tartaric Acid, Acetamide, L-Alaninamide, N-Acetyl-LGlutamic Acid, L-Arginine, Glycine, L-Histidine, L-Homoserine, Hydroxy-LProline, L-Isoleucine, L-Leucine, L-Lysine, L-Methionine, L-Ornithine, L-Phenylalanine, L-Pyroglutamic Acid, L-Valine, D,L-Carnitine, Sec-Butylamine, D.L-Octopamine, Putrescine, Dihydroxy Acetone, 2,3-Butanediol, 2,3-Butanone, and 3-Hydroxy 2- Butanone.

3.AX.3 Numerous ways not to isolate TM7.

Initial results from 454 pyrosequencing (not shown) of the third 7.2 $\times 10^6$ ppb isoprene sequential isoprene enrichment of willow soil yielded over 80% TM7 sequences. DGGE analysis (not shown) shown that the final communities were similar in each enrichment. As TM7 has also been enriched by stable isotope probing of soils with heavy benzene and toluene (Xie et al., 2011), this was interpreted at the time as an indication that TM7 may be involved in isoprene degradation. Following work, outlined in Chapter 2 shown this belief to be false, however by that time a multitude of techniques had been deployed to try an isolate isoprene-degrading TM7.

This section briefly describes the methods used in these unsuccessful attempts, as the mixture of tested and innovative methods lead to the isolation of a number of bacteria which may not have been isolated otherwise, and may help others in their attempts at isolating hard-to-culture species.

3.AX.3.1 Screening methods:

Two methods were used for screening for TM7 presence. The sequences from the two TM7 phylotypes acquired by 454 pyrosequencing were used to generate primer sequences specific to those isoprene-involved TM7.

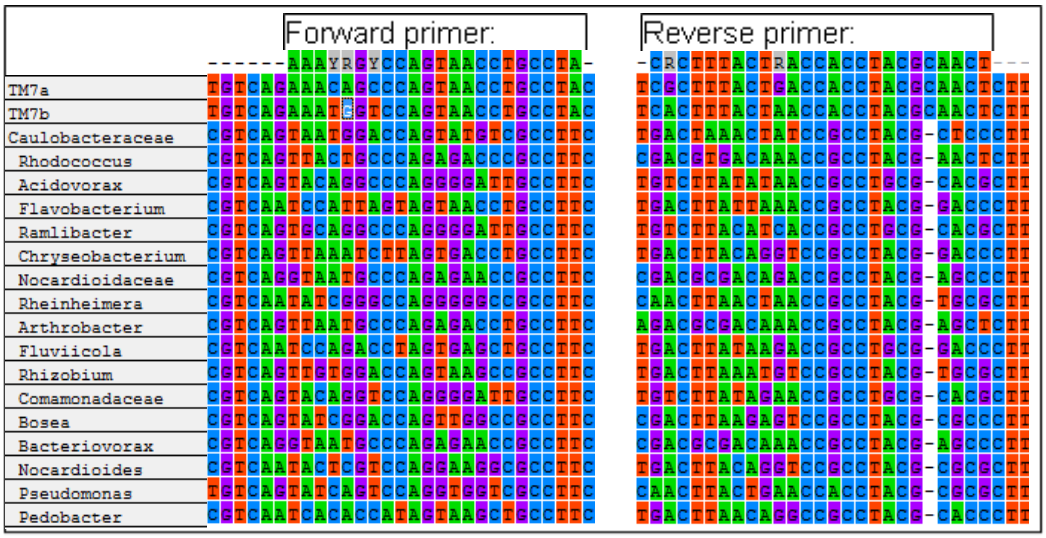


Figure 3.AX.3.1.1 Primer sequences with their target sequences of TM7 at the target area, alongside the same region for the 17 next abundant species in a TM7-rich isoprene enrichment of Willow soil.

Primer sequences were tested by BLASTn and both were specific. Oligonucleotide analysis was preformed using the Multiple primer analyzer by Thermo Scientific (http://www.thermoscientificbio.com/webtools/multipleprimer/), and found to be acceptable. A TM7 rich microcosm (10^4 dilution of the third generation of Willow soil) was amplified using these primers and gave the correct product size (175bp) and Sanger sequencing (Source Bioscience) yielded a product which when blasted was shown to be TM7, with negligible background noise (identical sequence to the TM7b from the pyrosequence data).

Thermocycling conditions were: 95˚C for 5 m, followed by 8 cycles of: 95˚C for 30 s, 60˚C, decreasing by 0.5˚C each time, for 30 s and 72˚C for 45 s. This was followed by 32 cycles of 95˚C for 30 s, 56˚C for 30 s, 72˚C for 45 s and ending with 72˚C for 10 m.

Products were ran on a 2% agarose gel in TAE buffer for 30 min at 110 V, stained with ethidium bromide and viewed under UV.

Colonies which were shown by PCR to contain TM7 were investigated for purity by Denaturing Gradient Gel Electrophoresis (DGGE).

DGGE was performed by mixing stock solutions of 0 % denaturant (20 ml Acrylamide, 2 ml of 50× TAE (242 g Tris, 57.1 ml glacial acetic acid and 100 ml of 500 mM EDTA per litre) and 78 ml dH_2O), 100 % denaturant (20 ml Acrylamide, 40 ml formamide, 2 ml of 50 × TAE, 42 g Urea) and 1% APS (0.1 g ammonium persulfate in 1 ml dH_2O).

A gradient gel was created between 60% denaturant (13.8 ml of 0% denaturant, 9.2 ml of 100 % denaturant) and 100% denaturant (9.2 ml 0% denaturant, 13.8 ml 100% denaturant) each with 90 µl APS and 9 µl TEMED (Tetramethylethylenediamine).

The gel was loaded and immersed in TAE and raised to 60°C and ran at 100 V for 1 hour and then 60 V overnight.

The gel was stained by fixing in fixing solution (100 ml ethanol, 5 ml acetic acid) for 30 min, stained in staining solution (0.2 g silver nitrate, 200 ml dH_2O, 1.6 ml formaldehyde) for 20 min, and developing solution (3 g NaOH, 200 ml dH_2O) 5-10 min, then fixing solution for 10 to 15 min.

The microcosms with the highest purity determined by visual inspection of DGGEs were used for the next stage of isolation in each method.

3.AX.3.2 Isolation methods

Based on the high level of TM7 in these cultures, the first methods attempted were through traditional cultivation on rich media (LB) agar plates, on minimal media (as in Chapter 2) agar plates, and through serial dilution of microcosms which were hoped to reduce diversity.

Following this, and the failure to isolate TM7, it was hypothesised that TM7, like some other recalcitrant bacteria (Vartoukian et al., 2010) may require a factor provided by another bacteria in order to survive. Cultivation attempts using this hypothesis were performed in several ways; Firstly, As we know *Rhodococcus* is also enriched, isoprene-degrading *Rhodococcus* isolates were grown up en-mass on agar plates, suspended in minimal media and lysed through bead beating. This lysate was then filter sterilised and applied either within the setting agar, or added to the agar surface. Secondly, as we know that the microcosm the TM7 is abundant in would have the dependencies of TM7, daughter microcosms were set up, lysed and the lysate was applied to isolation attempts. Thirdly as the soil the TM7 originated from was assumed to contain the requirements for TM7 growth, the soil was lysed, filtered and applied, and fourthly filters (<.22 µm, various materials) were superglued over circular incisions on foil, autoclaved, had a mixture of soil and minimal media applied to one side, and sealed, with the other side of the filter used as a cultivation surface (a budget approach based on Ferrari et al., (2005), who managed to achieve micro colonies of TM7 in a theoretically similar method). Fifthly, TM7

rich cultures were added to small, sterilised, 2 cm^3, GC-MS vials, with the centre of the rubber seal punched out, and a filter added in its place (sealed by compression), and were added to liquid microcosms in which TM7 rich microcosms were serially diluted. Sixthly, the TM7 rich microcosms was spread thinly onto minimal media plates, and then covered with two thin layers of minimal media with 3% washed agarose, which was then used as a cultivation surface (creating a barrier between the rich, and diluted microcosm). Seventhly, in an approach based on Stewart (2012), after spreading plates of TM7 rich microcosm, other isoprene-degrading isolates were dotted (live) onto the plate, which could have allowed for growth factor and other compound diffusion.

After no particular success through those methodologies, a number of competition inhibition based approaches were used to try and manipulate the diversity of the culture in a pro-TM7 way. In particular cultures with Streptomycin (10, 50, 100, 150 µg l^{-1}), Kanamycin (10, 50, 100, 150 µg l^{-1}) and EDTA (µg l^{-1}) additions were set up. Streptomycin resistance at the ribosome level is thought to be caused by a mutation (C to U at E. *Coli* 16S rRNA position 912) found in Archea and in TM7, but not in most other bacteria (Hugenholtz et al., 2001). This means that addition of streptomycin to TM7 containing microcosms, was hypothesised to increase the proportion of TM7, by inhibiting most other bacteria. Kanamycin targets primarily gram negative bacteria, and TM7 is believed to be gram positive (Hugenholtz et al., 2001), additionally EDTA has been shown to inhibit gram negative bacteria by removing Ca+ and Mg+ ions from phospholipids, causing them to repel each other, hence weakening the cell membrane, as well as inhibiting protazoa and decreasing

bacterial aggregation (Madden et al., 2004; Belfiore et al., 2007; Rupp et al., 1995).

Following this, as some phylotypes of TM7 have been shown by DNA-SIP experiments to degrade Toluene and Benzene (Luo et al., 2009; Xie et al., 2011), it was plausible that the TM7 phylotypes attempted to be isolated in this study had similar metabolic capabilities. Therefore microcosms with Toluene and Benzene additions (7.2×10^5 ppb), with and without Isoprene were set up, to try and increase abundance based on differential abilities to degrade other hydrocarbons.

Although TM7 is enriched significantly in minimal media microcosms with isoprene, it is unable to grow on minimal media agar with isoprene. The two obvious differences between the environments are the presence of agar, and the liquidity of the media. It was therefore speculated that TM7 required a liquid environment to thrive, or may be inhibited by agar, which could be due to either issues with attachment, inhibitory compounds or inability for dispersal of metabolic waste. As TM7 is able to grow in liquid minimal media, but not agar, the addition of liquid minimal media to plates would cause conditions to be more similar to those in liquid cultures, and may allow TM7 to grow, and also attach to the solid media, therefore liquid plates were set up with addition of 7 ml minimal media.

As growth is rich in minimal media, where it does not contact agar, solidifying agents other than agar, or non-agar attachment surfaces which allow nutrient diffusion will allow cultivation, therefore gelatin plates (5%), polyacrylamide plates (made with 4 ml acrylamide, 16 ml water, 3 µl APS and 3 µl TEMED. The

plate was soaked in distilled water overnight, drained, soaked in distilled water for 4 hours, drained, UV sterilised for 5 × 20 s (UVP CL-1000 Ultraviolet Crosslinker, max power), soaked in minimal media overnight, excess fluid was removed and the plate was inoculated and incubated in isoprene) and filter plates (made with PTFE, Cellulose Nitrate, Polycarbonate, RA, Cellulose Acetate, Nickel(II) oxide, Nylon, and Polycarbonate filters, of pore size <0.22 µm, which were laid onto cooling washed agarose minimal media) were created and analysed.

As the other possible difference between agar plates and media microcosms is the presence of glass, adding glass exposed beads may promote TM7 growth, or attachment to the glass. Due to this, beaded plates (beads (1.5 - 2 mm diameter) were dropped onto thin (~1 mm) washed agarose minimal media plates (3 per cm^2), with 5 mm minimal media layered on top, inoculated and incubated in isoprene overnight, drained and incubated with isoprene) were set up and analysed.

During testing using liquid plates, the other agar plates in the desiccation jars were found to have taken water aboard. This did not contaminate plates with low biomass (high dilutions) or blank controls, however the plates with growth exhibited higher growth than plates which were not incubated in the moist environment. As this accidently increased moisture led to increased growth from TM7 rich microcosms on minimal media agar plates, it was hypothesised that TM7 may require a more moist environment to grow, therefore artificially moist plates were created (with 1 ml minimal media additions to the lid).

TM7 rich microcosm floc formation and biofilm formation suggested the ability of TM7 to form flocs. These flocs are likely to be rich in EPS. TM7 may require EPS for growth, so addition of EPS to plates may allow colony growth. As EPS extraction methods destroy the structure of the EPS layers, the utilisation of EPS naturally produced by TM7 rich microcosms onto agar plates could be used instead. Therefore Biofilm plates were created (Liquid plates were inoculated and allowed to grow for two days, following which they were allowed to volatise for three hours, and drained. The agar plates had by that point created a light biofilm. The biofilm plates were UV sterilised for 7 × 20 seconds, inoculated with 10^2 and 10^4 dilutions of TM7 rich microcosms and incubated with isoprene alongside an un-inoculated control.

Unfortunately, all these methods were ultimately unsuccessful. From the SIP experiments (Chapter 2) we can now see why; TM7 is not growing on isoprene. Additionally from the metagenomic experiments, we can now generate a new hypothesis; that *Aeromicrobium* may be the species which TM7 requires to survive.

At the time Candidate division TM7 was un-isolated, and therefore isolation would have lead to an exciting insight into this large, ubiquitous, diverse and hard-to-study group, however since then a representative has been cultured (He et al., 2015), unfortunately the culture was not sustainable and has since been lost; and so there is still a lot to investigate about this group.

3.AX.3.3 Example image of gel showing PCR products:

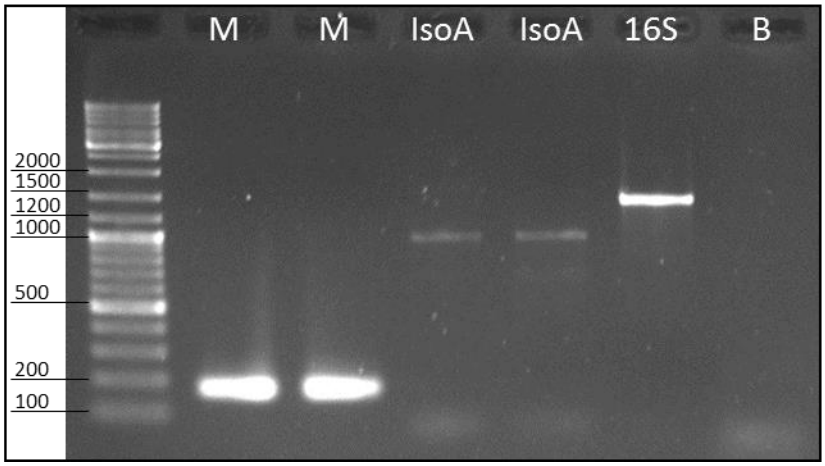


Figure 3.AX.3.4.1 Example PCR products for amplification of isolates using (M) Muyzer primers, (IsoA) *IsoA* primers, and (16S) 16S rRNA primers, with a 16S blank (B) ran on 1.5% agarose gel (80V 40 min), stained with ethidium bromide and displayed using UV light

3.AX.3.4 Appendix References:

Belfiore C, Castellano P and Vignolo G (2007) Reduction of *Escherichia coli* population following treatment with bacteriocins from lactic acid bacteria and chelators. *Food Microbiology.* **24**, 223-229.

Ferrari BC, Binnerup SJ and Gillings M (2005) Microcolony Cultivation on a Soil Substrate Membrane System Selects for Previously Uncultured Soil Bacteria. *Applied and Environmental Microbiology*. **71**, 8714-8720.

Hugenholtz P, Tyson GW, Webb RI, Wagner AM and Blackall LL (2001) Investigation of candidate division TM7, a recently recognized major lineage of the domain Bacteria with no known pure-culture representatives. *Applied and Environmental Microbiology.* **67**, 411–419.

Luo C, Xie S, Sun W, Xiangdong L and Cupples AM (2009) Identification of a Novel Toluene-Degrading Bacterium from the Candidate Phylum TM7, as Determined by DNA Stable Isotope Probing. *Applied and Environmental Microbiology*. **75**, 4644-4647.

Madden M, Song C, Tse K and Wong A (2004) The Inhibitory Effect of EDTA and Mg2+ on the Activity of NADH Dehydrogenase in Lysozyme Lysis. *Journal of Experimental Microbiology and Immunology*. **5**, 8-15.

Rupp DC, Anger C, Kapadia S and Totaro M (1995) Method Of Inhibiting Protozoan Growth In Eye Care Products Using A Polyvalent Cation Chelating Agent. Patent: US 5382599-A

Stewart EJ (2012) Growing Unculturable Bacteria. *Journal of Bacteriology*. **194**, 4151-4160.

Xie S, Sun W and Luo C (2011) Novel aerobic benzene degrading microorganisms identified in three soils by stable isotope probing. *Biodegradation*. **22**,71-81.