Performance evaluation of microbe and plant-mediated processes in phytoremediation of toluene in fractured bedrock using hybrid poplars

by

Michael Ben-Israel

A Thesis

presented to

The University of Guelph

In partial fulfilment of requirements
for the degree of

Doctor of Philosophy

in

Environmental Sciences

Guelph, Ontario, Canada

ABSTRACT

PERFORMANCE EVALUATION OF MICROBE AND PLANT-MEDIATED PROCESSES IN PHYTOREMEDIATION OF TOLUENE IN FRACTURED BEDROCK USING HYBRID POPLARS

Michael Ben-Israel
University of Guelph, 2020

Advisor:
Dr. Kari Dunfield

Efficacy of hybrid poplar trees for phytoremediation of toluene in fractured bedrock aquifers is unclear and active mechanisms require validation. This multi-year field study was conducted on a pilot phytoremediation system at an urban site, implemented to address aged toluene impacts to a shallow fractured dolostone aquifer. The study aimed to establish performance by quantifying phytoremedial activities at the site. Contaminant concentrations in groundwater, soil, and soil vapour, measured in high spatial resolution, showed the main residual toluene mass is coincident with the water table and located favourably for phyto-remedial uptake and biodegradation, with shallow groundwater concentrations approaching aqueous solubility in high-impact areas. Biodegradation occurring in the vadose zone was shown through metagenomic analyses that enumerated toluene degradation genes and gene transcripts in roots and root-associated soil, and compound-specific stable isotope analysis that showed enrichment of toluene stable carbon isotopes in soil vapour. Transpiration measurement, *in planta* contaminant quantification, and high-throughput sequencing of microbial taxonomic genes in roots and stem tissue were employed to measure toluene uptake through phytoextraction and resolve biodegradation influences upon uptake patterns. Though most phyto-available

toluene was being degraded/attenuated prior to uptake, phytoextraction rates were quantified in a subset of trees over a two-week peak-season period. Phytoextraction was greatest in the site's high-impact region. Trees there had distinct, more uniform root-colonizing bacterial communities, also surmised to have a greater toluene-degrading capacity compared to other locations. Stem phyllosphere microbiomes were shaped by *in planta* toluene presence as well, showing enrichment in predictive degradation capacity with increasing toluene exposure. Microbial diversity and richness in the phyllosphere were seasonally dynamic, increasing in the late growing season. Finally, a lab-scale DNA stable isotope probing study identified putatively novel toluene-degrading bacteria and fungi taxa in rhizosphere soil and their taxonomic gene sequences were made available for future studies. This study validated ways in which phytoremediation of toluene using hybrid poplars actively occurs in fractured bedrock systems and resolved contributing chemical and biological mechanisms of action on quantitative and qualitative scales. Techniques employed in this study broaden available field-validated methods to monitor and assess aromatic hydrocarbon attenuation in poplar phytoremediation systems.

DEDICATION

To Judy.

ACKNOWLEDGEMENTS

First, thank-you to my advisor, Dr. Kari Dunfield, for providing me with this unique opportunity, for your mentorship, and for your steadfast support and guidance throughout this journey. I would also like to acknowledge and thank the other members of my advisory committee, Dr. Beth Parker, Dr. Elizabeth Haack, and Dr. David Tsao, for your time, guidance, and indispensable advice. I am grateful to my entire committee for this immersive, practical experience across many fields and disciplines.

I would like to acknowledge the other members of the collaborative research and development program supported by BP, the University Consortium for Field-Focused Groundwater Contamination Research, and the Natural Sciences and Engineering Research Council of Canada (NSERC). Thank-you to Dr. Ramon Aravena for your guidance to help keep me on track. Thank-you to Jeremy Fernandes, for initiating and onboarding me into this project and for the solid foundation your work provided for my research. I am grateful also to Dr. Philipp Wanner, for your exceptional collaboration, helpful discussions, and technical support. Thank-you to Dr. Joel Burken, who's expertise advanced the breadth of this work. Thank-you for technical support, training, and field supervision provided by partners at AECOM and EcoMetrix Inc., including Jeff McBride, Sean Todd, Matt Smith, Dr. Fei Luo, and many others. A special thank-you to Alan Scheibner (and colleagues at BP Canada) for your unwavering support to this project and your commitment to scientific development and education.

I am grateful for invaluable technical support, training, and helpful discussion from members of the Dunfield research group, G^{360} research group, and the University community at large. A special thank-you to Kamini Khosla, for your time and contributions. Thank-you to Jonathan Gaiero, Dr. Jemaneh Habtewold, Dr. Micaela Tosi, Tolulope Mafa-Attoye, Anibal Castillo, Dr. Dasiel Obregon Alvarez, John Drummelsmith, Andrea Roebuck, Dr. Nicola Linton, Travis Mazurek, Dr. Elizabeth Bent, Dr. Eduardo Kovalski Mitter, Sara Low, Adrianna Wiley, James Hommersen, Maria Gorecka, Isaac Noyes, Steve Chapman, Juliana Camillo, Dr. Patryk Quinn, Rashmi Jadeja, Amanda Pierce, Sean Jordan, Ian Renaud, Steve Wilson, Dr. Dyanne Brewer, Dr. Armen Charchoglyan, Dr. James Longstaffe, Brent Coleman, and Kevin Ecott among many others.

I am forever grateful for the constant support I have received from my family and friends, without whom I would not be where I am today. Finally, my utmost gratitude goes to Pooja Arora – thanks for inspiring me every day and pushing me to achieve my goals.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ags	Above ground surface
ANOVA	Analysis of variance
BD	Buoyant density
bgs	Below ground surface
bssA	Benzyl succinate synthase alpha subunit
BTEX	Benzene, ethylbenzene, toluene, xylene(s)
CSIA	Compound-specific isotope analysis
DFN	Discrete Fracture Network
DGGE	Denaturing gradient gel electrophoresis
DNA-SIP	DNA stable isotope probing
GC-IRMS	Gas chromatograph coupled to an isotope mass spectrometer
GC-MS	Gas chromatograph coupled to a mass spectrometer
HGU	Hydrogeological unit
LNAPL	Light non-aqueous phase liquid
PBS	Phosphate-buffered saline
PCoA	Principal coordinates analysis
PERMANOVA	Permutational analysis of variance
PHE	Phenol hydroxylase
PICRUSt2	Phylogenetic Investigation of Communities by Reconstruction of Unobserved States 2
qPCR	Quantitative real-time PCR
RMO	Ring-hydroxylating monooxygenase
SPME-GC	Solid phase microextraction gas chromatography
SPS	Solid phase sampler
ST1–3	Sampling trenches 1, 2, and 3
SV	Sequence variant
SVP1–3	Soil vapour probes 1, 2, and 3
TDP	Thermal dissipation probe
TOD	Toluene dioxygenase
VPDB	Vienna Pee Dee Belemnite
WMS	Waterloo Membrane Sampler

1 Introduction

1.1 Motivation, study site, and research program

Performance evaluation is a vital element of environmental contaminant remediation programs to meet regulatory guidelines and site management goals. This thesis presents an evaluation of an established remediation technology, phytoremediation using hybrid poplar trees, to address subsurface impacts of a monoaromatic hydrocarbon, toluene, in a complex hydrogeological formation where its mechanistic efficacy was unclear. Toluene is part of the benzene, toluene, ethylbenzene, and xylene isomer (BTEX) family of volatile organic compounds. Toluene is stored in pure form and in mixtures in high volumes globally. It occurs naturally as a constituent of crude oil and petroleum products and is used broadly in industry as a carrier solvent for commercial products such as paints, adhesives, and pharmaceuticals (OECD 2012). Terrestrial and aquatic releases of toluene occur as light non-aqueous-phase liquids (LNAPLs), typically from leaking underground storage tanks, seepage from landfill sites, municipal and industrial waste discharge, and through accidental spills (CCME 2004). When release volumes are sufficiently large, they produce persistent contamination in groundwater and soil. Release of toluene and other BTEX compounds into soil and groundwater poses a risk to human and environmental health and safety and must be addressed for appropriate risk management and regulatory compliance (WHO 2004; MOECP 2011; OECD 2012).

The study site for this research was a former manufacturing facility operational between 1952 and 1990 in a mixed industrial-residential neighbourhood in South-Western Ontario, Canada. During operation, it historically housed an outdoor toluene storage tank and buried distribution line that supplied toluene to the facility. Upon decommissioning of the tank and product line in 1989, evidence of toluene impacts to the fractured sedimentary dolostone bedrock aquifer beneath were discovered and groundwater monitoring was initiated. Traditional site characterization, hydraulic containment, and

remediation methods were attempted in the following decades but were discontinued due to ineffectiveness and other site-specific concerns within the urban-residential setting. In 2008, a hybrid poplar (*Populus deltoides* × *nigra* OP-367) phytoremediation system of 51 trees was planted above the former storage tank and supply line for remediation and containment of the toluene source and dissolved-phase plume (Figure 1.1).

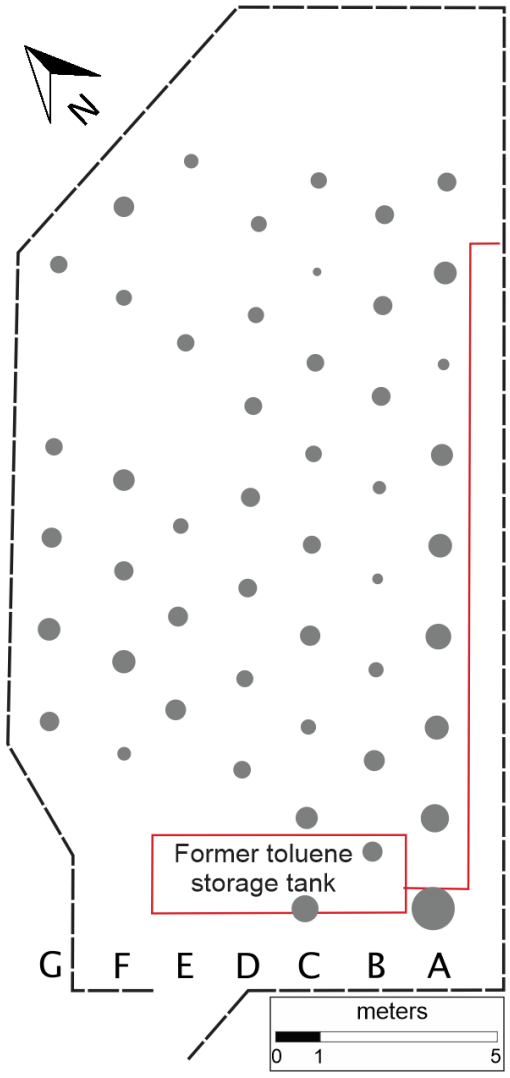


Figure 1.1: Fenced hybrid poplar phytoremediation area showing trees (grey circles reflecting relative trunk sizes *ca*. 2016) in rows A–G over the former locations of the toluene storage tank and buried distribution line (red lines).

Phytoremediation using hybrid poplar trees is part of a broader set of phytotechnologies, defined as "*technologies using plants to remediate or contain contaminants in soil, groundwater, surface water, or sediments*" (ITRC 2009b). These methodologies are particularly suited in urban and remote environments, where infrastructure and issues of public perception may favour passive treatment options. For aromatic hydrocarbons and other volatile organics, shallow occurrence in the saturated and vadose zones makes them amenable to phytoremediation by plants, such as poplar trees, whose roots can persist in saturated conditions (Ferro et al. 2013). The main mechanisms of action include contaminant biodegradation by root-associated microorganisms (rhizodegradation) and uptake and removal through the transpiration stream (phytoextraction and phytovolatilization) (Jordahl et al. 1997; Burken and Schnoor 1998; Limmer and Burken 2016). Hybrid poplars are facultative phreatophytes (can access water from the saturated zone), therefore hydraulic control over a groundwater plume is possible also, depending on site-specific hydrology and tree water use rates (Ferro et al. 2001). Where root interaction with impacted media is achieved, rhizodegradation constitutes the dominant mass reduction end-fate in these systems (Limmer et al. 2018). Over long-term applications, phyto-enhanced biodegradation can decrease dissolved-phase concentrations of contaminants in groundwater, lowering the demand on oxygen supplied through seasonal recharge, and thus shift aquifer redox status over time (Landmeyer and Effinger 2016). Often, the dominant attenuation mechanisms for toluene phytoremediation applications are microbe-mediated (Weishaar et al. 2009; Limmer et al. 2018). The genetic capacity for toluene biodegradation is relatively well understood for bacteria and some fungi and occurs through several aerobic and anaerobic pathways (Figure 1.2; Parales et al. 2008). In aerobic pathways, bacteria will initiate degradation through hydroxylation reactions using mono- and dioxygenase enzymes targeting the toluene aromatic ring or methyl side-chain. These pathways converge on production of catechol and catechol derivatives, which are then cleaved by dioxygenase enzymes and converted into Krebs cycle metabolites. Anaerobically,

bacteria will initiate degradation through fumarate addition to the methyl side-chain, then similarly converge on production of Krebs cycle metabolites. Marker genes in these degradation pathways, typically those encoding degradation-initiating enzymes, can be detected in remediation systems using molecular biological tools in order to evaluate remedial capability and assess community ecological patterns (Baldwin et al. 2003; Winderl et al. 2007; Nebe et al. 2009).

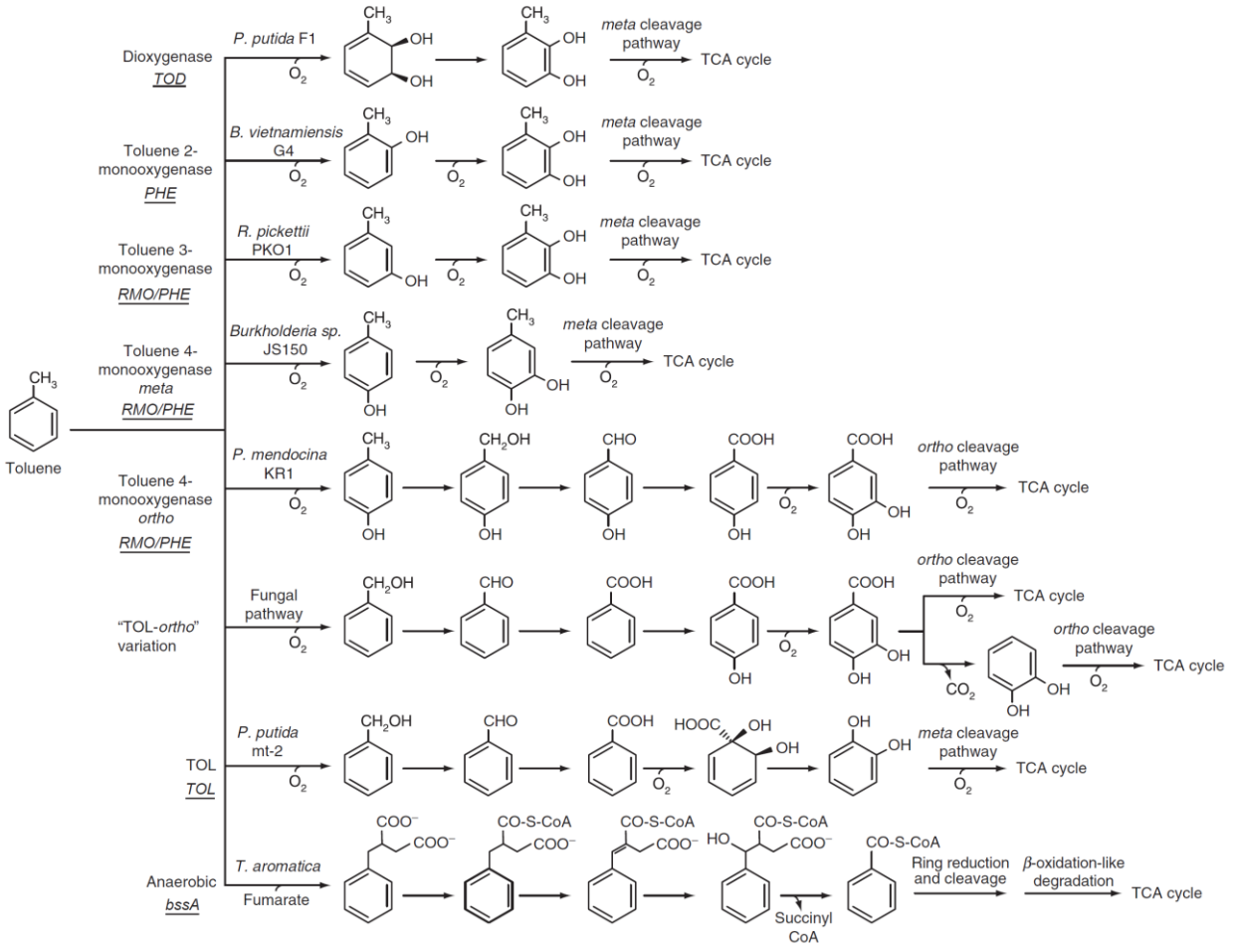


Figure 1.2: Key bacterial and fungal toluene biodegradation pathways; modified from Parales et al. (2008). Common PCR primers for aerobic TOD, PHE, and RMO (Baldwin et al. 2003) and anaerobic *bssA* (Winderl et al. 2007) gene targets are indicated in italics/underlining under

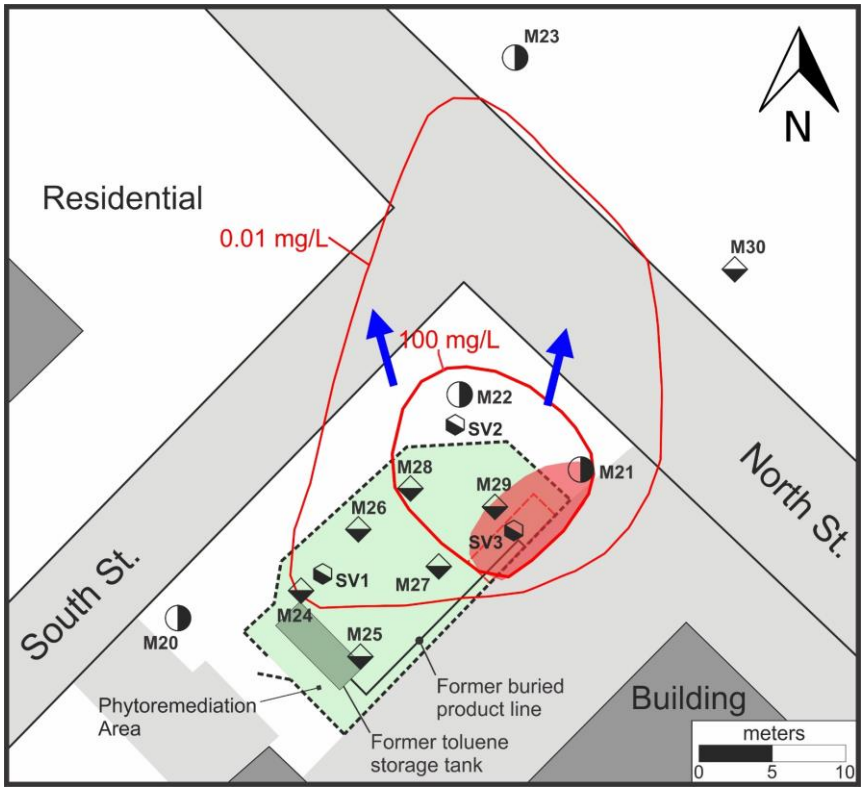
associated pathway names. After initial oxidation, the aromatic ring is cleaved and converted into TCA (Krebs) cycle metabolites and incorporated into central metabolism.

Phytoremediation applications to-date have shown success at attenuating subsurface organic impacts under conditions where groundwater is present in unconsolidated materials (El-Gendy et al. 2009; Ferro et al. 2013; Landmeyer and Effinger 2016). For the present site, occurrence of toluene as LNAPL in fractured rock created circumstances where efficacy is unclear. LNAPL behaves differently in fractured rock compared to unconsolidated and single-porosity consolidated sediments. They can experience greater penetration into the saturated zone due to large pressure heads of LNAPL entering vertical fractures, fluctuations of the water table can increase lateral migration as LNAPL is pushed through fractures, and LNAPL sorption and diffusion into bedrock matrix can produce retarded source areas that sustain persistent groundwater plumes (Hardisty et al. 2004; Parker et al. 2012). As such, unique contaminant mass and phase distribution and heterogeneous groundwater flow paths through fracture networks make contaminant interaction with plants unclear, thus the efficacy of phytoremediation may be decreased. Conversely, roots from shallow-growing trees have been observed to extend along and into bedrock fractures to access groundwater (Zwieniecki and Newton 1995; Hubbert et al. 2001); therefore, it was theorized that poplar root growth at this site might advance into the fracture network and thereby target remediation of impacted media.

A multidisciplined collaborative research program was established in 2014, when the trees had reached maturity, to characterize the site in sufficient detail regarding groundwater flow regimes and contaminant mass, phase distribution in relevant media (bedrock, groundwater, soil, and vapour) to allow for conclusions to be made for the efficacy of the phytoremediation program. The collaborative team included researchers from the University of Guelph, the G^{360} Institute for Groundwater Research, the University of Waterloo, Missouri University of Science and Technology, and industry partner BP and

its environmental consultants. The goals of the program included compilation of a highly refined conceptual site model of toluene concentrations and mass distribution within the fractured bedrock and fracture network, delineation of the dissolved-phase toluene impacts in groundwater, and refined understanding of natural toluene (bio)attenuation mechanisms in the saturated and vadose zones. Since then, a site conceptual model developed using the Discrete Fracture Network (DFN) approach (Parker et al. 2012) identified that toluene mass and phase distribution on site is variable across the planted area, while the majority of mass is found in shallow bedrock, coincident with the fluctuating water table, and is thus located favourably for phytoremediation (Figure 1.3; Fernandes 2017). Comprehensive efforts at site characterization and the conceptual site model are outside the scope of, but are referred to where relevant in this thesis, which focused on elucidating plant-associated biodegradation of toluene and quantifying its uptake by the poplar stand. From a contaminant mass-balance perspective, assessing phytoremediation efficacy *in situ* is difficult due the superimposition of plant attenuation mechanisms on the fate and transport of the contaminant in the subsurface and the dynamic nature of the system. An interdisciplinary multiple-lines-of-evidence approach utilizing ecological assessment of microbial degrader populations and novel, cutting-edge contaminant monitoring techniques was central to this work to evaluate efficacy.

A)



B)

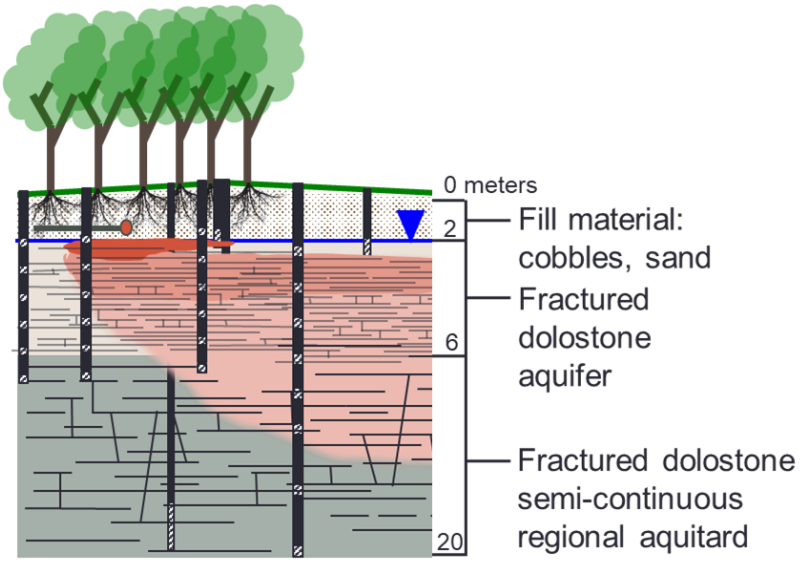


Figure 1.3: Conceptual model of contaminant distribution on site during beginning of phytoremediation research program; modified from Fernandes (2017). (A) Site map showing toluene groundwater source area (red shading), groundwater plume extent (red circles), site multi-

level well and soil vapour sampling infrastructure (half-shaded diamonds, circles, and hexagons), and fenced phytoremediation area. (B) Vertical cross-section showing depth of overburden, mean water table position at approximately 2 m below ground surface (blue line with triangle), approximate depths of site multi-level wells, and the shallow occurrence of the bulk of toluene mass near the bedrock-overburden contact (red shading).

1.2 Thesis format and objectives

This thesis is presented in manuscript-style chapters that are published (Chapters 2 and 3) or are in preparation for submission to peer-reviewed journals (Chapters 4 and 5). The overall objectives of this work were (1) to determine quantitatively and mechanistically if phytoremediation using hybrid poplars can influence attenuation of toluene from a source in a fractured bedrock aquifer system; and (2) to further develop field-validated methodologies to study and characterize remediation processes *in situ*. Specific objectives of the research were addressed in the individual research chapters:

Chapter 2 presents a study assessing active toluene biodegradation in the vadose zone through multiple lines of evidence;

Objective 2.1: establish evidence of active toluene biodegradation in the vadose zone;

Objective 2.2: characterize degrader abundances and ecology in the rhizosphere and root microbiome;

Chapter 3 explores the relationship between toluene phytoextraction and the influence of biodegradation on uptake and detection;

Objective 3.1: quantify active toluene phytoextraction with site-wide spatial resolution;

Objective 3.2: identify root-colonizing bacterial communities and their functional relation to detected phytoextraction profiles;

Chapter 4 expands the investigation to microbial communities within above-ground plant tissue and how toluene exposure shapes their structure;

Objective 4.1: characterize stem phyllosphere endophyte populations seasonally;

Objective 4.2: quantify toluene-degrading functional capacity in relation to toluene exposure;

Chapter 5 broadens the scope of known toluene degraders through targeted assessment of DNA from toluene-utilizing microbes;

Objective 5.1: isolate DNA specifically from toluene-degrading bacteria and fungi;

Objective 5.2: identify toluene degraders, including potentially novel organisms;

The final chapter of this thesis (Chapter 6) summarizes the findings of Chapters 2 to 5 and discusses their implications for site phytoremediation and monitoring.

2 Toluene biodegradation in the vadose zone of a poplar phytoremediation system identified using metagenomics and toluene-specific stable carbon isotope analysis

This manuscript has been published in the International Journal of Phytoremediation:

BenIsrael M, Wanner P, Aravena R, Parker BL, Haack EA, Tsao DT, Dunfield KE. 2019. Toluene biodegradation in the vadose zone of a poplar phytoremediation system identified using metagenomics and toluene-specific stable carbon isotope analysis. Int J Phytoremediation. 21(1):60-69. https://doi.org/10.1080/15226514.2018.1523873

Keywords: BTEX; CSIA; gene expression; qPCR; rhizosphere; root microbiome

2.1 Abstract

Biodegradation is an important mechanism of action of phytoremediation systems, but performance evaluation is challenging. We applied metagenomic molecular approaches and compound-specific stable carbon isotope analysis to assess biodegradation of toluene in the vadose zone at an urban pilot field system where hybrid poplars were planted to remediate legacy impacts to an underlying shallow fractured bedrock aquifer. Carbon isotope ratios were compared spatio-temporally between toluene dissolved in groundwater and in the vapor phase. Enrichment of ^{13}C in toluene in the vapor phase compared to groundwater provided evidence for biodegradation in the vadose zone. Total bacterial abundance (16S *rRNA*) and abundance and expression of degradation genes were determined in rhizosphere soil (DNA and RNA) and roots (DNA) using quantitative PCR. Relative abundances of degraders in the rhizosphere were on average higher at greater depths, except for enrichment of PHE-encoding communities

that more strongly followed patterns of toluene concentrations detected. Quantification of RMO and PHE gene transcripts supported observations of active aerobic toluene degradation. Finally, spatially-variable numbers of toluene degraders were detected in poplar roots. We present multiple lines of evidence for biodegradation in the vadose zone at this site, contributing to our understanding of mechanisms of action of the phytoremediation system.

2.2 Introduction

Non-halogenated petroleum hydrocarbons, such as benzene, toluene, ethylbenzene, and xylene (BTEX), are common subsurface contaminants at numerous commercial gas station and industrial sites worldwide (Holliger et al. 1997), and are often released into the subsurface as lighter-than-water, non-aqueous phase liquids (LNAPLs). If release volumes are sufficiently large, LNAPLs can reach the water table, where they float and spread laterally due to their density being less than water, and can result in persistent contaminant plumes (ITRC 2009a). Due to their high vapor pressure, LNAPLs also volatilize and migrate upward into the vadose zone forming vapor phase contaminant plumes (Kim and Corapcioglu 2003; Christophersen et al. 2005). The shallow occurrence of these contaminants in the saturated and vadose zones makes them amenable to phytoremediation (ITRC 2009b). Hybrid poplar trees have been demonstrated to be an effective application of this phytotechnology, being able to contain and even fully remediate petroleum hydrocarbon-contaminated sites (Ferro et al. 2001; Barac et al. 2009; El-Gendy et al. 2009; Landmeyer and Effinger 2016). Plant-associated bacteria play an integral role in remedial activity, especially in the vadose zone, where transpiration draws contaminants and oxygen into the rhizosphere and into contact with degrader populations (Weishaar et al. 2009). However, evaluating activity and overall phytoremediation efficacy is challenging and requires well-informed, effective monitoring strategies.

Metagenomic molecular techniques are powerful emerging tools that can be used to delineate the distribution and activity of marker degradation genes across dynamic field conditions in order to understand the ecology, capacity, and activity of degrader organisms in remediation systems (Larentis et al. 2013). Metabolism of petroleum hydrocarbons by degraders as carbon and energy sources can produce shifts in community composition; thus, differences in proportions of toluene degraders relative to total bacterial abundance across contamination gradients can serve as a line of evidence for biodegradation (Bell et al. 2013). Tracking of degradation marker genes provides a means for evaluating remedial capability (Winderl et al. 2008; Brow et al. 2013; Key et al. 2014) and elucidating complex ecological patterns of communities in applied remediation systems (Nebe et al. 2009; Baldwin et al. 2010; Lunsmann et al. 2016b). However, to reliably overcome limitations in assay specificity and environmental variability, molecular techniques should be supported by additional lines of evidence (Smith and Osborn 2009; Iwai et al. 2011). Compound-specific isotope analysis (CSIA) is as an effective tool that can be combined with molecular techniques to offer significant insight into *in situ* biodegradation processes affecting organic contaminants (Illman and Alvarez 2009). To detect biodegradation, the CSIA method makes use of the preferential cleavage of bonds between light compared to heavy isotopes that leads to a continuous enrichment of heavy isotopes in the parent compared to the daughter compound (Braeckevelt et al. 2012). Initially, CSIA was mainly applied in groundwater systems; however, recent studies have demonstrated its applicability also in the vadose zone (Bouchard et al. 2008a; Bouchard et al. 2008b; Goli et al. 2017). Molecular techniques have been employed previously in combination with CSIA in both field and lab-scale investigations (Beller et al. 2008; Rakoczy et al. 2011), but have not yet been employed in tandem to study *in situ* biodegradation in the vadose zone of phytoremediation systems.

To-date, examination of plant-associated biodegradation within poplar-based petroleum hydrocarbon phytoremediation systems has often been conducted on the lab-scale, limited to culture-dependent methods (Jordahl et al. 1997; Barac et al. 2009;

Weishaar et al. 2009; Timm et al. 2015), or involved uncontaminated field conditions (Ulrich et al. 2008; Gottel et al. 2011; Beckers et al. 2017). The root microbiome of poplars grown under field conditions is known to harbor a distinct assemblage of organisms compared to associated rhizosphere as a result of environmental differences, rather than opportunistic colonization of roots by dominant rhizosphere organisms (Gottel et al. 2011); however, environmental selective pressures are poorly defined. Information is lacking also on the ecological differences between degraders associated with the rhizosphere, poplar root surfaces (rhizoplane), and within the root microbiome *in situ*. These critical knowledge gaps are significant for evaluating biodegradative processes and are relevant also for evaluating other phytoremediation activities, such as mass removal through phytoextraction (Wilson et al. 2013; Limmer et al. 2018).

In the present study, we assessed the efficacy of a small-scale (220 m^2 planted area) poplar phytoremediation system accessing toluene mass from soil and a shallow, transient, fractured bedrock flow system. Quantitative real-time PCR (qPCR), reverse transcription-qPCR, and CSIA were employed to investigate biodegradation processes affecting toluene in the vadose zone, which operate in conjunction with plant mass-removal processes. Herein, we provide results relating to plant-microbe interactions and their effects on abundance and expression of degradation genes, and on isotopic signatures of biodegradation within the vadose zone. The present study provides a greater understanding, supported through multiple lines of evidence, of the ecology and activities of bacterial toluene degrader populations in phytoremediation systems that is paramount for effective corrective action planning, implementation, and site monitoring.

2.3 Materials and Methods

2.3.1 Site description

The investigated site was a manufacturing facility in a mixed residential/industrial neighborhood in urban South-Western Ontario, Canada. Toluene was used as a solvent

in manufacturing processes and was stored in an outdoor storage tank. Toluene impacts to soil, bedrock, and groundwater were discovered in 1989 during decommissioning of the partially buried tank and a buried product distribution line (Figure 2.1A).

Site geology includes a 2-m-thick overburden layer of sandy, cobbled soil with remnants of former concrete infrastructure, overlying a 21-m-thick sequence of Silurian-era, fractured dolostone bedrock. A suite of high-resolution field methodologies, collectively known as the Discrete Fracture Network (DFN) approach (Parker et al. 2012), was recently applied to characterize site hydrogeology and the distribution and phase of toluene in the fractured rock and groundwater (Fernandes 2017). Eleven multi-level wells (six to eight monitoring ports) were installed to a maximum depth of 22 m below ground surface (bgs). Three hydrogeological units (HGUs) were delineated based on vertical hydraulic head profiles. The shallowest HGU, spanning 2-6.5 m bgs, shows a strong vertical hydraulic gradient, and low vertical connectivity to underlying HGUs. The water table fluctuates between 2.1 and 3.0 m bgs, coincident with the bedrock-overburden contact, and groundwater flows in a northern direction showing a velocity of <0.01 – 3.0 m/day (Figure 2.1A).

A primary groundwater source zone of residual toluene was identified in the north-eastern corner of the site (Figure 2.1A), wherein toluene occurs as residual LNAPL, and in the dissolved and sorbed phases within the shallow bedrock. Dissolution of LNAPL into groundwater and back-diffusion from the low-permeability dolostone bedrock matrix is sustaining a persistent plume (Figure 2.1A). Approximately 95% of dissolved-phase toluene is found within the upper two meters of the saturated zone, corresponding to the water table depth in deep overburden during high water table conditions, and uppermost bedrock. As such, the main residual toluene mass is located favorably for phyto-remedial uptake and/or biodegradation. Residual impacts to soil have not previously been characterized at this site.

A phytoremediation tree stand of 51 hybrid poplars (*Populus deltoides* × *nigra* OP-367) was installed in 2008 in a dense formation overlying the former tank and supply line (Figure 2.1A). Mean tree diameter at breast height was approximately 15.97 cm in September 2016 prior to the time of this study. Goals of the phytoremediation application included plume containment and uptake and remediation of toluene in impacted soil, groundwater, and bedrock. Trees were planted in nutrient-amended boreholes to promote rapid establishment and deep rooting; boreholes were advanced to refusal up to a maximal depth of approximately 2 m bgs at the bedrock-overburden contact. This approach has been shown to increase interaction with impacted media and speed up remedial and containment activities (Ferro 2003).

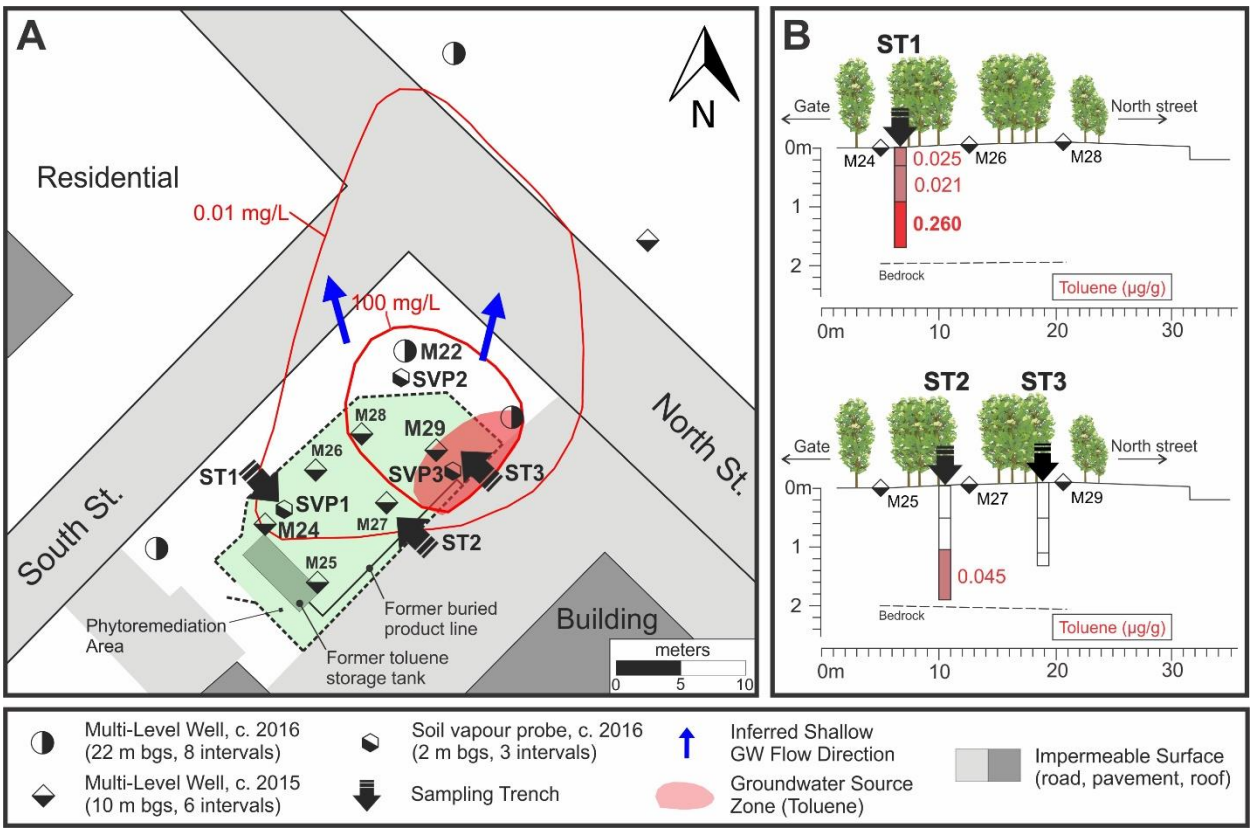


Figure 2.1: Investigation site map showing toluene groundwater source and plume extent, site infrastructure, and phytoremediation area (A), and 2016 sampling trench cross-sections with

measured soil toluene concentrations (B). Hybrid poplar trees were planted in a 220 m^2 area c. 2008. Site infrastructure was installed c. 2015-2016 and used to infer toluene groundwater source and plume extent, and for collection of groundwater and soil vapor for toluene concentration and stable carbon isotope analyses. Sampling trenches were advanced over three days in November 2016 for collection of rhizosphere soil, rhizoplane soil, and root samples for molecular analysis.

2.3.2 Soil vapor and groundwater sampling

Soil vapor for toluene concentration analysis and CSIA was sampled from three soil vapor probes at the site (SVP1, SVP2, and SVP3; Figure 2.1A). For concentration analysis, soil vapor was collected monthly between July 11 and September 18, 2017 with 1-L Tedlar® bags using an SKC XR500 air pump at a pump rate of 450 ml/min with dedicated 0.64 OD Teflon® tubing. For CSIA, soil vapor was sampled passively with Waterloo Membrane Samplers (WMSs) (Seethapathy and Gorecki 2011). The WMSs were installed for 69 days between July 11 and September 18, 2017. This period was necessary to accumulate enough contaminant mass for CSIA. Groundwater for toluene analysis and CSIA was sampled monthly between July 11 and September 18, 2017, coinciding with soil vapor sampling for concentration analysis, from adjacent multilevel wells (M22, M24, and M29; Figure 2.1A). Groundwater was sampled from the shallowest port of the multilevel wells adjacent to the soil vapor probes using Geopump peristaltic pumps with dedicated 0.32 cm OD diameter Teflon® tubing. It was not possible to sample the shallowest port of multilevel M29 on August 13 and September 18, 2017 due to the low water table, therefore the second-shallowest port was sampled instead.

2.3.3 Toluene concentration analysis and compound-specific carbon isotope analysis (CSIA)

Detailed descriptions of toluene concentration analysis and CSIA can be found in Appendix A. Briefly, toluene concentrations in the vapor phase and in groundwater were analyzed using a gas chromatograph coupled to a mass spectrometer (GC-MS).

Compound-specific stable carbon isotope ratios of toluene were determined by a gas chromatograph coupled to an isotope mass spectrometer (GC-IRMS). Carbon isotope signatures were analyzed relative to the Vienna Pee Dee Belemnite (VPDB) standard and results were expressed in the delta notation: δ_{VPDB} (‰) = $(R/R_{std}-1)\times1000$, where R and R_{std} are the isotope ratios of the sample and the standard, respectively. The measured $\delta^{13}C$ values in the vadose zone were corrected by adding 1‰ to account for isotope fractionation during sampling, caused by diffusion through the WMS membrane and sorption in the WMS (Goli et al. 2017).

2.3.4 Rhizosphere soil, rhizoplane soil, and root samples

Poplar roots and soil samples were collected over three days in November 2016 from sampling trenches excavated at three locations, upgradient of (ST1 and ST2) and above (ST3) the toluene groundwater source zone (Figure 2.1). Weather conditions were reflective of very mild autumn with ambient temperatures of approximately 13-18 °C, the ground was not frozen, and trees had not entered senescence. Mechanical advancement of single trenches, versus multiple trenches in a given area, was necessitated by large cobbles and/or buried infrastructure in the overburden material. Trenches were advanced using a CAT® 420 excavator with an 18-inch bucket in three depth intervals until refusal at final depths of 1.70, 1.93, and 1.37 m bgs (Table A.1).

Root samples with adhering soil (*ca*. 18 g) were collected from each depth (n=3), then separated into roots and rhizosphere soil by shaking off and collecting adhering soil. Subsamples of soil and roots were flash-frozen in the field by submerging in liquid nitrogen for 3 min, then transported on dry ice and stored at -80 °C until further processing. Additional rhizosphere soil subsamples were analyzed for toluene at Maxxam Analytics (Mississauga, Canada) according to EPA 8260C protocols, and gravimetric moisture content was determined within 48 h.

Rhizoplane samples were obtained as follows in the lab: *ca*. 6 g of root material was added to 250 mL sterile 1x phosphate-buffered saline (PBS), vortexed briefly, shaken at 220 rpm at 10 °C for 20 min, then rinsed in 100 mL sterile 1x PBS. Wet soil was recovered from the solution by centrifugation at 13,700×g for 20 min. Wet soil and surface-cleaned roots were immediately used for DNA extraction to analyze rhizoplane soil and root microbiome bacterial communities, respectively.

2.3.5 Extraction of nucleic acids and cDNA synthesis

The quantity and purity of all extracted DNA and RNA were assessed using a NanoDrop™ 8000 Spectrophotometer (Thermo Fisher Scientific Corp.). Total DNA and RNA were co-extracted from rhizosphere soil (*ca*. 1.4 g) using an RNA PowerSoil® Total RNA Isolation Kit and DNA Elution Accessory Kit (Mobio Laboratories Inc.) according to the manufacturer's instructions. Extracted DNA for use in qPCR was stored at -20 °C until analysis. The RNA extracts were immediately DNase-treated using RQ1 RNase-Free DNase (Promega Corp.) in quadruplicate reactions according to the manufacturer's instructions. The RNA was immediately converted to cDNA using a High Capacity cDNA Reverse Transcription Kit (Applied Biosystems Corp.) according to the manufacturer's instructions using 1 µL of SUPERase•In™ RNase Inhibitor (Life Technologies Corp.). DNase treatment efficacy was evaluated by visualizing DNase-treated and untreated RNA on agarose gels and by qPCR analysis.

Surface-cleaned roots (*ca*. 3 g) were homogenized using an MM400 Mixer Mill (Retsch GmbH) by agitating flash-frozen grinding jars twice at 25 Hz for 20 s. Total root DNA was extracted from root homogenate (0.3 g each in two pooled extractions) using a PowerSoil® DNA Isolation Kit (Mobio Laboratories Inc.), and gravimetric moisture content of root homogenate was determined.

Total rhizoplane soil DNA was extracted using a PowerSoil® DNA Isolation Kit (Mobio Laboratories Inc.) with modification to the manufacturer's protocol. DNA was

extracted following kit instructions for wet soil samples; centrifugation was increased to 3 min, and the process was repeated until achieving *ca*. 0.4 g of pelleted material. Pellet weights were later used for standardizing rhizoplane gene quantities.

2.3.6 Quantitative PCR

Total bacteria and toluene-degrading bacterial populations were assessed by quantifying genes and transcripts (cDNA) in the metagenomic samples using previously published primers. The 335f/769r primer set was used to quantify the V3-V4 regions of the bacterial 16S rRNA gene to estimate total bacterial abundance (Dorn-In et al. 2015). Aerobic degradation pathways were targeted using toluene dioxygenase (TOD), ring-hydroxylating monooxygenase (RMO), and phenol hydroxylase (PHE) primers (Baldwin et al. 2003). Anaerobic pathways were targeted using the 7772f/8543r primer set that target benzyl succinate synthase alpha subunit (*bssA*)-related genes encoding fumarate-adding enzymes (von Netzer et al. 2013). All qPCR reactions were carried out on a CFX96™ Real-Time PCR Detection System (Bio-Rad Laboratories Inc.) and conformed to MIQE guidelines (Bustin et al. 2009). Full details on primers used and reaction conditions can be found in Appendix A and Table A.2.

Gene and transcript values were normalized to grams dry soil (rhizosphere soil), dry root, or grams of pelleted extractant (rhizoplane soil) to relay results on a biologically significant scale. Negligible differences in DNA extraction efficiency between samples were assumed. Relative toluene degrader abundances were calculated by standardizing degradation gene quantities to 16S rRNA gene quantities in order to determine if changes in toluene degradation gene abundance reflected changes in bacterial abundance or shifts in community composition.

2.3.7 Statistical analyses

Statistical analyses (ANOVA) were performed to evaluate fixed effects and to compare means of gene and transcript quantities and relative gene abundances using SigmaPlot (v. 12.5) software. Significant differences in means were determined using the Holm-Šídák multiple comparisons test. Where necessary, Box-Cox power transformations ($\lambda = -2$ or 0) were performed prior to analysis to ensure normal distribution and equal variance of data. Significance was determine using a threshold of $\alpha=0.05$.

2.4 Results and discussion

2.4.1 Toluene concentrations in soil, groundwater, and soil vapor

Toluene measurements in rhizosphere soil showed a spatial distribution of mass between depths and locations across the site (Figure 2.1B). Toluene was detected in soil from all three sampled depths of ST1 (0.025-0.26 $\mu g/g_{soil}$). In ST2, toluene was only detected in the deepest interval (0.045 $\mu g/g_{soil}$). No toluene was detected in soil obtained from ST3, near the groundwater source zone. During trench excavation at ST3, excavation depth was limited to 33-56 cm shallower than the other locations due to buried infrastructure, possibly limiting our ability to sample soil at the groundwater interface, where we expect toluene levels to be highest.

Toluene concentrations in groundwater in the shallowest port of the multilevel wells showed minor temporal variation over the study period (Figure 2.2). The lowest concentrations in groundwater (0.79-15.52 $\mu g/L_{water}$) were measured in the inferred upgradient well (M24; Figure 2.2A). In contrast, concentrations were elevated (94,850-323,200 $\mu g/L_{water}$) in the inferred downgradient direction (M22; Figure 2.2B) and in the source zone (M29; Figure 2.2C). When the uppermost port at M29 was dry, toluene concentrations measured in the deeper port were somewhat lower (2,683-13,800

$\mu g/L_{water}$), consistent with the bulk of toluene mass residing at the approximate depth of the water table (Figure 2.2C).

Toluene was also detected in vapor collected from all three soil vapor probes installed in the vadose zone, at concentrations significantly lower than in groundwater (0.00279-0.111 $\mu g/L_{vapor}$; Figure 2.2). The lowest toluene concentrations in vapor over the study period were typically detected under seasonally high water table conditions (Figure 2.2). This suggests a potential relationship between exposure of the high toluene concentration zone in the shallow bedrock and vapor concentrations in the vadose zone that merits further investigation.

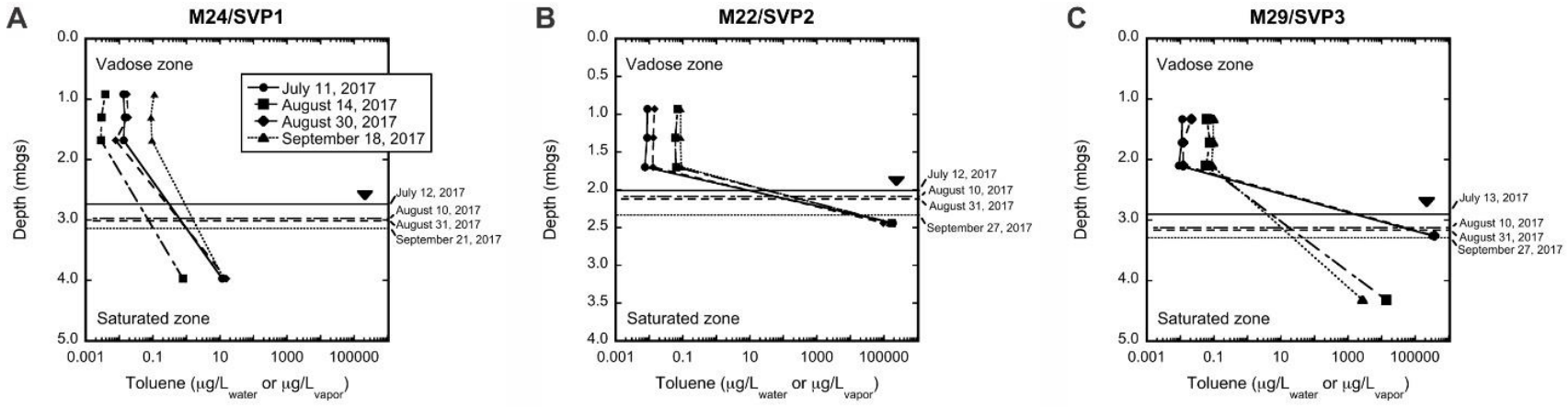


Figure 2.2: Measured toluene concentrations in soil vapor probes 1 (A), 2 (B), and 3 (C) in $\mu g/L_{vapor}$ (upper three data points), and in the shallowest port of the adjacent multilevel wells (M24, M22, and M29) in $\mu g/L_{water}$ (lowest data point) between July 11 and September 18, 2017. Horizontal lines indicate groundwater levels on sampling dates. The shallowest port in M29 was dry on August 14 and September 18, 2017 due to the low water table, therefore toluene concentrations from the second-shallowest port are presented instead (C). Note that toluene concentrations are plotted on a logarithmic scale to facilitate comparison between groundwater and vapor phase.

2.4.2 Stable carbon isotope ratios of toluene

Carbon isotope signatures of toluene in groundwater exhibited small temporal variations during the sampling period (Figure 2.3). The $\delta^{13}C$ values adjacent to the source zone (M29) were in a range between -29.82‰ and -29.26‰ and can be used as an approximation for the carbon isotope composition of the toluene groundwater source (Figure 2.3C). The $\delta^{13}C$ signatures in the upgradient and downgradient multilevel wells (M22 and M24) were consistent with signatures at M29, despite variation in toluene concentration between those locations.

In the vadose zone at the upgradient location (SVP1), a depletion of ^{13}C in toluene, i.e. a more negative $\delta^{13}C$ value, was observed just above the water table compared to the groundwater ($\Delta\delta^{13}C$ = 1.6‰), whereas at shallower depths the $\delta^{13}C$ values were similar to the groundwater (Figure 3A). In contrast, a less negative $\delta^{13}C$ value and thus an enrichment of ^{13}C in toluene was observed in the vadose zone compared to the groundwater at the downgradient location ($\Delta\delta^{13}C$ = 2.1‰) and at the location close to the groundwater source zone ($\Delta\delta^{13}C$ = 1.9‰), where high toluene concentrations were detected (Figure 3B and 3C). For a system without biodegradation, a depletion of ^{13}C equaling a more negative $\delta^{13}C$ value is expected in the vadose compared to the saturated zone, as isotopically light molecules diffuse faster than heavy molecules (Wanner and Hunkeler 2015). Thus, observed enrichment of ^{13}C in the vadose zone downgradient and adjacent to the source zone in comparison to groundwater provides one line of evidence for microbial toluene degradation in the vadose zone (Bouchard et al. 2008a; Bouchard et al. 2008b). Molecular tools targeting bacterial toluene degradation genes and transcripts were applied to soil samples to provide additional lines of evidence for toluene biodegradation within the vadose zone at this site.

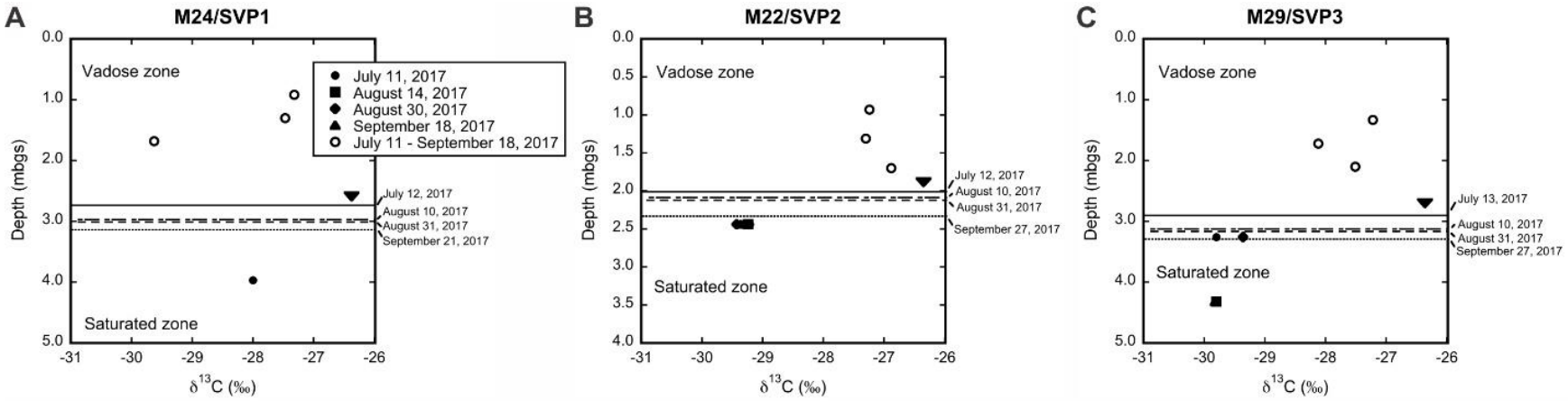


Figure 2.3: Measured stable carbon isotope ratios of toluene in soil vapor probes 1 (A), 2 (B), and 3 (C), and in the shallowest port of the adjacent multilevel wells (M24, M22, and M29) between July 11 and September 18, 2017. Horizontal lines indicate groundwater levels on sampling dates.

2.4.3 Abundance of toluene degradation genes in poplar-associated rhizosphere soil

Toluene degraders equipped with the genetic capacity for aerobic and anaerobic toluene biodegradation were detected in all rhizosphere soil sampled across the planted area and gradient of detected toluene concentrations in rhizosphere soil (Table A.1). Biodegradation genes were detected in the range of 10^5-10^6 (RMO and PHE), 10^4-10^5 (TOD), and 10^6 (*bssA*) copies per gram dry soil. Detected quantities of 16S rRNA genes (10^7-10^9 copies per gram dry soil) revealed that total bacterial abundance decreased significantly with increasing depth in ST1 and ST3 (Table A.1). Relative abundances of degradation genes, as proportions of total bacteria, reflected differences in bacterial community composition across the site, and indicated enrichment of bacterial populations capable of both aerobic and anaerobic toluene degradation at greater depths and toluene concentrations (Figure 2.4). The effects of sample location and depth on enrichment were evaluated; a significant increase in relative abundances of all toluene degradation genes was observed with increasing depth. Average relative abundances increased between depths 1 and 3 by 0.18% (RMO), 0.13% (PHE), 0.01% (TOD), and 0.82% (*bssA*) (Figure 2.4). Relative abundances were not significantly different between locations for TOD and RMO targets (Figure 2.4B&C) but changed significantly with depth for *bssA* and PHE targets (Figure 2.4A&D). Most notably, PHE gene enrichment was significantly greater in depth 3 of ST1 and ST2 (where toluene was highest; Figure 2.1B) compared to ST3.

Degrader enrichment is a known marker for contaminant biodegradation (Lovley 2003). Therefore, the overall enrichment of toluene-degrading populations relative to total bacterial numbers and the patterns described above provide a line of evidence for toluene biodegradation in the vadose zone. Toluene is found in vapor phase across the vadose zone (Figure 2.2); thus, evidence of biodegradation in rhizosphere soil in which toluene was not detected was not unexpected (Hohener et al. 2006). Biodegradation is likely influencing community composition more strongly at depth, where nutrients are more

limited (Jobbágy and Jackson 2001) and where soil toluene concentrations were typically observed to be higher, thus giving organisms with the capability to biodegrade toluene a competitive advantage. Increased abundance of degrader organisms in poplar rhizospheres has previously been shown to be a result of a general rhizosphere effect of the trees on contaminant-selected soil communities (Jordahl et al. 1997), and enriched degraders in the rhizosphere have been found to drop below detection once contaminant sources are fully removed or remediated (Barac et al. 2009). Therefore, we do not attribute degrader enrichment in rhizosphere soil at this site to specific biochemical and ecological plant-microbe interactions that are involved in other phytoremediation systems (Siciliano et al. 2002; Thijs et al. 2016).

Enrichment of PHE genes most strongly followed patterns of toluene concentrations detected in rhizosphere soil (Figure 2.4D). This suggests that organisms equipped with toluene biodegradation pathways that use PHE-encoded enzymes are dominating in locations within the vadose zone where toluene concentrations in rhizosphere soil are highest. Toluene monooxygenase genes phylogenetically related to the *tbmD* gene of *Burkholderia* sp. strain JS150, which are amplified by PHE but not RMO primers, may be responsible for observed enrichment of these populations (Nebe et al. 2009). We observed enrichment of organisms having anaerobic *bssA* genes (Figure 2.4A). These results do not, however, necessarily suggest that biodegradation was substantially occurring in the vadose zone under anoxic conditions. Increased oxygen infiltration into the vadose zone is typical in poplar phytoremediation systems due to transpiration activity and resultant depression of the water table (Weishaar et al. 2009). We observed a seasonal flux in the height of the water table in this manner below the poplar stand (Figure 2.1 and Figure 2.2). Moreover, subsequent preliminary sampling of the three soil vapor probes has demonstrated that the vadose zone exists under oxic conditions (data not shown). It is therefore possible that facultative anaerobic toluene degraders possessing both anaerobic and aerobic toluene degradation pathways, such as *Thauera* sp. strain DNT-1 (Shinoda et al. 2004), were being enriched.

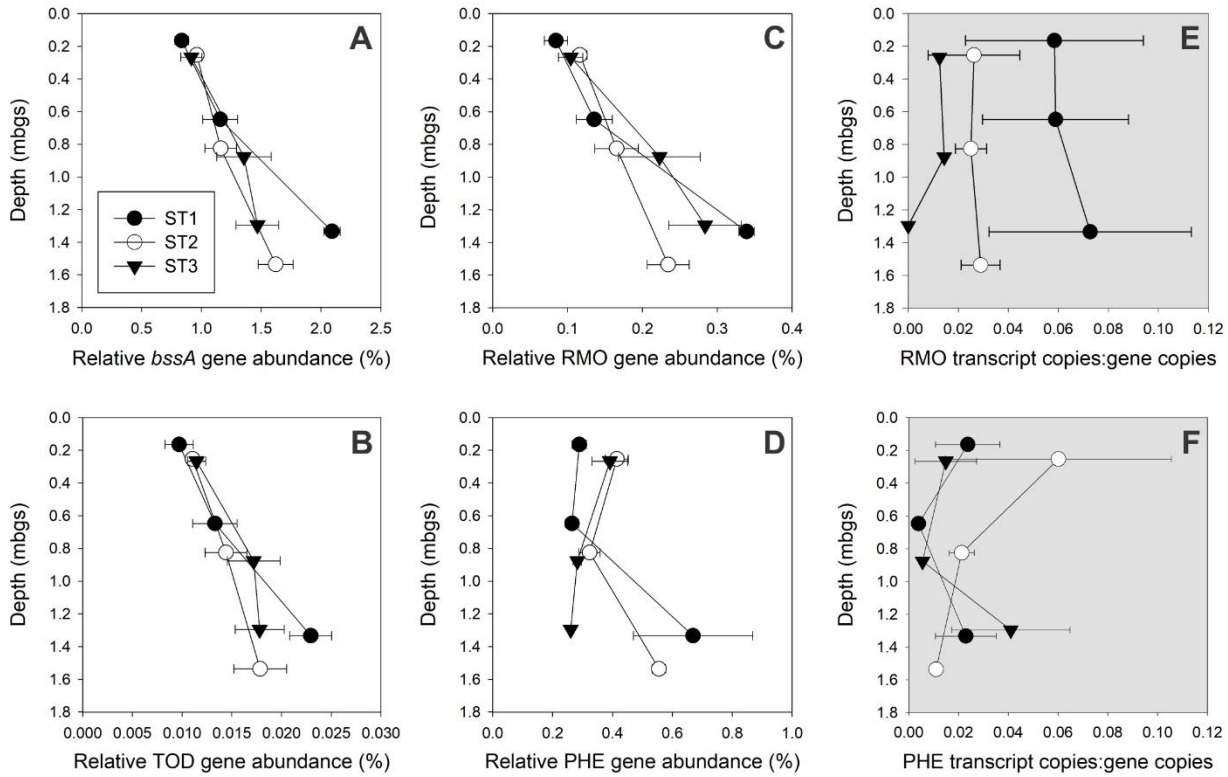


Figure 2.4: Relative gene abundances of toluene-degrading bacteria (A-D), and ratios of aerobic transcripts relative to genes (E and F) in rhizosphere soil. Relative gene abundances as proportions of bacterial 16S rRNA genes (%) in rhizosphere soil for anaerobic *bssA* (A) and aerobic TOD (B), RMO (C), and PHE (D) genes. Aerobic transcripts as ratios to gene copies for RMO (E) and PHE (F). Error bars are standard error of replicates. For relative gene abundances, n = 3, except ST3 depth 3, where n = 2. For transcript to gene ratios, n = 3, except RMO (ST1 depth 2, ST2 depth 1, and ST3 depth 1) and PHE (ST3 depth 1), where n = 2.

2.4.4 Toluene degradation gene expression in poplar-associated rhizosphere soil

Active expression of degradation gene targets in rhizosphere soil was evaluated by quantifying mRNA transcripts of RMO and PHE gene targets to determine whether these genes were simply present, or indeed actively being utilized. Aerobic toluene biodegradation genes were actively being transcribed at the time of sampling, reinforcing that biodegradation is occurring in the vadose zone of the phytoremediation system

(Table A.1). Transcript quantities ranged from non-detect to 10^4 (RMO) and 10^2-10^5 (PHE) copies per gram dry soil. Transcripts were detected at all locations, except for RMO in depth 3 of ST3, where genes were potentially being transcribed at levels below our limit of detection. On average, RMO transcripts differed significantly between locations, and were more abundant in ST1 and ST2, where rhizosphere soil toluene concentrations were highest. However, PHE transcripts did not follow the same trend, despite PHE gene enrichment at these locations (Figure 2.4D). Expression of BTEX degradation genes has previously been shown in contaminated soils to be highest at locations with the greatest availability of substrate and/or electron acceptor (Key et al. 2014). However, degradation gene expression is reflective of metabolic processes that are transient in nature, with numerous fluctuating environmental factors, such as nutrient and oxygen status, pH, moisture, and temperature affecting associated transcript quantities (Saleh-Lakha et al. 2005). Therefore, expression patterns may be reflecting conditions at the time of sampling where metabolic activity and therefore gene induction were low. Relative transcript quantities (as proportions of corresponding gene quantities) were evaluated in an attempt to identify zones where high-level expression might be occurring (Brow et al. 2013), but did not show distinct patterns (Figure 2.4E&F). Induction of degradation gene expression occurs in the presence of toluene; but terminal electron acceptors are occasionally sufficient inducers, even in the absence of toluene (Brow et al. 2013). Moreover, both aerobic and anaerobic genes can be expressed at low levels in preparation for changing environmental conditions that favor biodegradation, such as temperature, oxygen status, and substrate concentration (Shinoda et al. 2004). Oxygen concentrations were not measured at the time of sampling, but oxygen has been detected throughout the vadose zone in preliminary sampling of the three soil vapor probes (data not shown). Thus, PHE gene targets may have had sufficient inducers of expression available across the site, while induction of RMO targets may be more dependent on access to toluene. Together, these results provide evidence of a dynamic response of bacterial toluene degrader populations in the vadose zone to environmental conditions and toluene concentrations

and support the finding that enrichment of degraders in the rhizosphere is a result of biodegradation.

2.4.5 Abundance of toluene degradation genes in poplar roots and root-associated rhizoplane soil

Toluene degradation gene abundances (RMO, TOD, and PHE) in rhizoplane soil and root microbiomes indicated that degrader communities changed across the soil-root interface at this site (Table A.3). Quantities of 16S rRNA genes ranged from 10^8-10^9 copies per gram wet rhizoplane soil and 10^9-10^{10} copies per gram dry root. Degradation gene quantities per gram wet rhizoplane soil ranged from 10^5-10^6 (RMO), 10^4-10^5 (TOD), and 10^5-10^7 (PHE). Gene copies per gram dry root ranged from 10^6-10^7 (RMO) and 10^5-10^6 (TOD). PHE genes were present in all root samples (10^6-10^7); however, quantification was limited by a secondary PCR product (~100 bp) that amplified in some samples. Off-target plant DNA can sometimes interfere with metagenomic study of bacteria in the root microbiome (Arenz et al. 2015); therefore, specificity of PHE primers merits further examination when looking at root-associated communities. Degrader abundances did not differ significantly between roots where toluene was detected in associated rhizosphere soil compared to roots with rhizosphere soil having no detected toluene (data are presented in the Supporting Information). All roots are considered toluene-exposed due to vapor-phase toluene in the vadose zone (Figure 2.2), and the small-scale nature of the site precluded access to non-exposed control roots.

The relative abundance of RMO and TOD genes to 16S rRNA genes (Figure 2.5) showed changes in the composition of toluene degrader populations associated with poplar roots. Relative degradation gene abundance changed significantly with averages of 0.304% (RMO) and 0.0165% (TOD) in rhizoplane soil, and 0.471% (RMO) and 0.0217% (TOD) in roots. Relative TOD abundance differed by sampling location, illustrating significantly different enrichment patterns between and within rhizoplane soil and roots (Figure 2.5B&E). A greater relative abundance of TOD genes in roots in depth

3 was observed in ST3 compared to ST1 and ST2 (Figure 2.5E). These results suggest that distinct degrader assemblages are populating poplar root microbiomes across the site, and that degraders are more active and thus being enriched in deeper roots of trees overlying more heavily-impacted groundwater in the toluene groundwater source zone. Given that poplars at this site are of the same genotype and age, access to toluene is likely the dominant selective pressure on community composition in their root microbiomes (Ulrich et al. 2008). Efforts to measure *in planta* concentrations of toluene and degrader abundances in above-ground tissues along transpiration streams of trees at this site are currently underway. Wilson et al. (2013) previously showed that biodegradation prior to plant uptake directly impacts the concentrations of BTEX in poplar transpiration streams; however, the influence of biodegradative processes along the transpiration stream has remained relatively unexplored and merits further investigation.

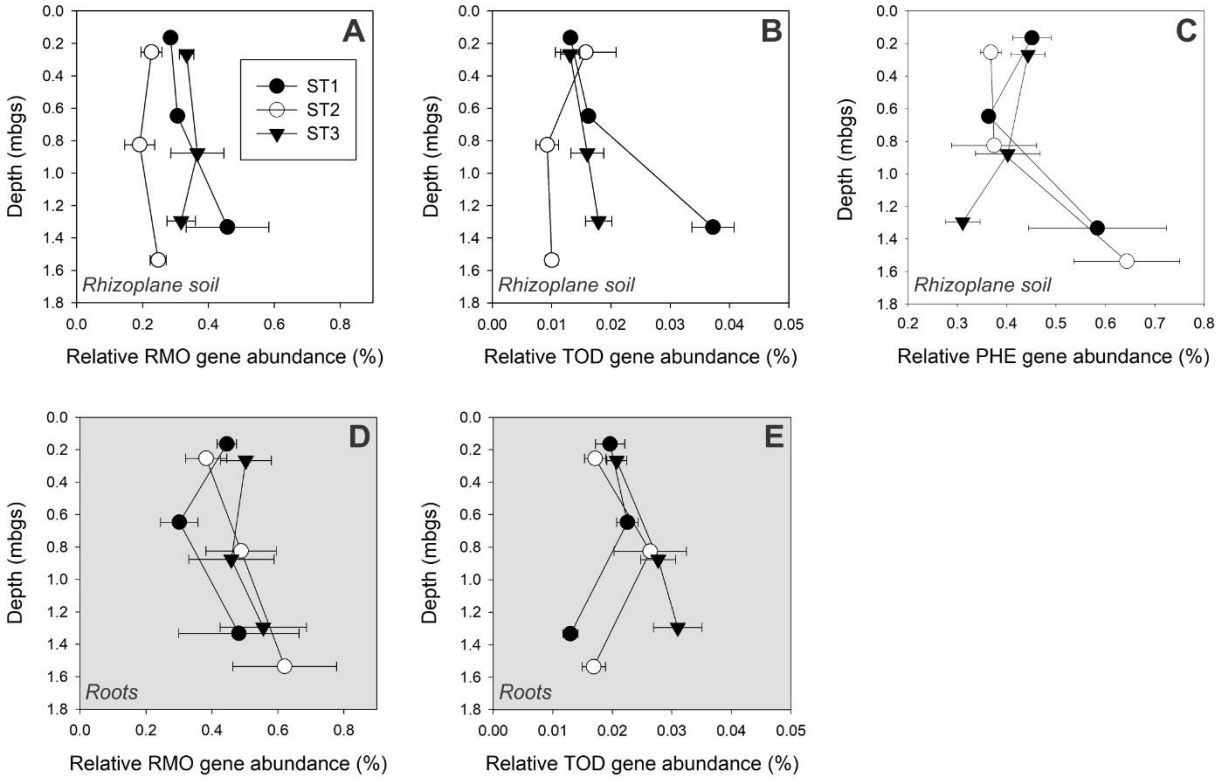


Figure 2.5: Relative gene abundances of toluene-degrading bacteria associated with poplar rhizoplane soil (A-C) and roots (D and E). Relative gene abundances as proportions of bacterial 16S rRNA genes (%) in rhizoplane soil and roots for RMO (A, D) and TOD (B, E). Note, PHE relative abundances in roots are not reported due to co-amplification of non-specific targets in root samples. Error bars are standard error of replicates (n = 3, except rhizoplane soil from ST1 depth 3, where n = 2).

2.5 Conclusions

In this study, we applied molecular and stable isotope methods as lines of evidence supporting biodegradation of toluene in the vadose zone of a 9-year-old poplar phytoremediation system established over a shallow, fractured bedrock aquifer contaminated with toluene. Our findings corroborate previous evidence in the literature that a general rhizodeposition effect influences toluene degradation gene abundance in the rhizosphere, and that access to contaminants strongly affects community

composition. Our results indicate that organisms inhabiting poplar rhizosphere and root microbiome compartments play a role in biodegradation in the vadose zone. Finally, we provide evidence herein suggesting that tree proximity to impacted groundwater may be shaping degrader community structure in poplar roots at greater depths, and we propose that strategies monitoring phytoextraction of petroleum hydrocarbons should also investigate degradative capacity of endophytic communities along transpiration streams. Together, degrader enrichment in rhizosphere and roots highlight how phytoremediation systems couple contaminant uptake with enhanced biodegradation for broader remedial activity. This research provides improved knowledge of the ecological structures of poplar-associated degraders and will aid in future efforts to enhance phytoremediation application and monitoring strategies, and to quantify processes to inform remedial time-scales.

Acknowledgements:

This work was supported by BP, the University Consortium for Field-Focused Groundwater Contamination Research, and the Natural Sciences and Engineering Research Council of Canada (NSERC) through a Collaborative Research Development Grant. We thank Alan Scheibner (BP Canada) for ongoing support and contributions to the project.

3 Quantification of toluene phytoextraction rates and microbial biodegradation functional profiles at a fractured bedrock phytoremediation site

This manuscript has been published in Science of the Total Environment:

BenIsrael M, Wanner P, Jeremy Fernandes J, Burken J, Aravena R, Parker B, Haack E, Tsao D, Dunfield K. 2020. Quantification of toluene phytoextraction rates and microbial biodegradation functional profiles at a fractured bedrock phytoremediation site. Sci Total Environ. 707: 135890. https://doi.org/10.1016/j.scitotenv.2019.135890

Keywords: BTEX hydrocarbons; endophyte; molecular biological tools; root microbiome; site remediation; site monitoring

3.1 Abstract

This field study evaluated the efficacy of a mature hybrid poplar phytoremediation system for the remediation of toluene in a fractured bedrock aquifer site. Phytoextraction activity of the trees and the ecology and biodegradation potential of root-colonizing bacteria that ultimately influence how much toluene is transported from the roots and phytoextracted to the aboveground point of measurement were explored. Peak-season toluene mass removal rates ranging from 313 to 743 µg/day were quantified using passive *in planta* contaminant sampling techniques and continuous heat dissipation transpiration measurements in tree stems. Root bacterial microbiome structure and biodegradation potential were evaluated via high-throughput sequencing and predictive metagenomic functional modelling of bacterial 16S *rRNA* genes in roots. Poplar roots were colonized mostly by Proteobacteria, Actinobacteria, and Bacteroidetes. Distinct, more uniform communities were observed in roots associated with trees planted in the toluene source

area compared to other areas, with differences apparent at lower taxonomic levels. Significant enrichment of *Streptomyces* in roots was observed in the source area, implicating that genus as a potentially important poplar endophyte at toluene-impacted sites. Moreover, significantly greater aerobic toluene biodegradation capacity was predicted in these roots compared to other areas using taxonomic functional modelling. Together with passive sampling, the molecular results provided supporting evidence of biodegradation activity in the source area and contextualized the detected phytoextraction patterns. These results support the application of phytoremediation systems for aromatic hydrocarbons in environments with complex geology and demonstrate field-validated monitoring techniques to assess phytoextraction and biodegradation in these systems.

3.2 Introduction

Effective risk management measures, bolstered by reliable monitoring techniques, are essential to address anthropogenic releases of aromatic hydrocarbons, such as benzene, toluene, ethylbenzene, and xylene (BTEX), into the environment (Apitz et al. 2005; Karr and Chu 2008). Moderate water solubility and volatility of these compounds make them permissibly mobile in the subsurface, acute and prolonged exposure pose considerable health and safety risks, and releases of substantial volume, therefore, can pose a significant threat to drinking water resources (WHO 2004; OECD 2012). Phytoremediation, defined as the use of plants for attenuation of environmental contaminants through remediation and containment, is an increasingly popular, passive risk management strategy for impacted sites (Schwitzguébel et al. 2002; Thijs et al. 2017). Depending on site management time-frames and goals, this technology can be considered suitable and economically favorable so long as immediate contaminant exposure risks are low (Compernolle et al. 2012). Economic considerations of phytoremediation are reviewed further in Gatliff et al. (2016).

Phytotechnologies have successfully been employed using hybrid poplar (*Populus spp.*) trees for remediation and risk management of BTEX-impacted sites (El-Gendy et al. 2009; Landmeyer and Effinger 2016). These trees augment natural attenuation mechanisms in the vadose and saturated zones, including through hydraulic capture/containment of groundwater plumes (Ferro et al. 2001; Barac et al. 2009) and enrichment and enhancement of soil and root-associated degrader microbial communities (Jordahl et al. 1997; Bourdel et al. 2016). Full mineralization is achieved by BTEX-degrading bacteria; therefore, no secondary pollutants are expected to be generated in these applications (Parales et al. 2008). Contaminant removal can be achieved also through phytoextraction and phytovolatilization – i.e. root uptake and translocation of contaminants from soil, water, or vapor, followed by atmospheric release (Burken and Schnoor 1999; Limmer and Burken 2016; Wilson et al. 2017; Limmer et al. 2018). As such, no contaminant accumulation is expected in these systems. In phytoextraction of organic compounds, uptake is passive with the transpiration stream and should not be confused with phytoextraction of inorganic contaminants, which involves root cell-mediated active transport mechanisms (ITRC 2009b). On a seasonal basis, uptake has been observed to be highest in the summer months, when transpiration rates are highest (Landmeyer and Effinger 2016).

Phytoremediation applications to-date have shown success under conditions where groundwater is present in unconsolidated materials (El-Gendy et al. 2009; Ferro et al. 2013; Landmeyer and Effinger 2016). However, applications are limited and efficacy is unclear for systems addressing impacts to fractured bedrock aquifers, which represent substantial drinking water sources that are vulnerable to contamination (Palau et al. 2014; Allen et al. 2017). Differences in behavior and storage of BTEX releases occur in fractured rock compared to unconsolidated sediments due, in part, to light, non-aqueous phase liquid (LNAPL) fluid dynamics within fracture network geometries, their sorption and dissolution within rock matrices, and heterogenous groundwater flow paths in these systems (Hardisty et al. 2004). As a result, unique contaminant mass, phase distribution

and groundwater flow through fracture networks might add complexity to the interactions of plant roots with contaminants. However, trees planted into shallow soils overlaying fractured bedrock have been shown to extend roots along and potentially into fractures to access stored groundwater (Zwieniecki and Newton 1995; Hubbert et al. 2001). Thus, the potential for phytoremediation systems to be uniquely poised to effectively remediate residual impacts in fractured bedrock systems needs to be explored.

Of key importance in establishing phytoremediation performance is demonstration and quantification of phytoextraction activity. Quantification of *in situ BTEX* phytoextraction rates has been limited to-date, and involves quantifying water use and *in planta* contaminant concentrations (Landmeyer and Effinger 2016; Limmer et al. 2018). Heat dissipation methods are often used in estimating tree transpiration rates (Ferro et al. 2001; Zalesny et al. 2006), and can effectively be employed to assess whole-stand water use over time by extrapolating single-tree estimates over entire tree stands using an estimate of the basal trunk area measured at standard breast height (Kostner et al. 1998; Wullschleger et al. 1998). Contaminant uptake and *in planta* concentrations can be evaluated using a number of tree sap and tissue analysis methods, but these are often poorly suited for trace-level detection or for capturing temporally variable transpiration or contaminant concentrations within source media (Limmer et al. 2014a; Limmer et al. 2014b; Limmer and Burken 2015). A simple, cost-effective alternative has been found using passive *in planta* solid phase samplers (SPSs), which provide time-weighed contaminant concentrations on more relevant temporal scales (Shetty et al. 2014). These samplers are also more suitable for long-term, repeat analysis because they do not require repeat damage to trees. It should be noted that when quantifying *in planta* BTEX constituents, detected concentrations do not always reflect subsurface contaminant concentration profiles (Wilson et al. 2013). This is due to their high rates of biodegradation and attenuation of mass that occurs prior to plant uptake (Weishaar et al. 2009) and along the transpiration pathway by internal plant tissue-colonizing endophyte microbes (Barac

et al. 2004; Barac et al. 2009). Understanding the spatio-temporal nature and extent of biodegradation is, therefore, essential for contextualizing other phytoremedial activities.

Measurement of *in situ* biodegradation rates at impacted field sites remains particularly elusive (Illman and Alvarez 2009), though exploration of inferred relative activity is possible. Taxonomic metagenomic techniques can be used to assess *in situ* microbes on a whole-community scale and to infer degrader activities based on ecological structure (Moore et al. 2006; Nolvak et al. 2012; Douglas et al. 2018). Furthermore, emerging functional evolutionary modeling methods, such as PICRUSt2, can also be employed to predict metabolic functional capacity in metagenomic samples by relating detected taxonomic marker genes, such as 16S *rRNA* genes, to reference genome databases (Langille et al. 2013). These taxonomic functional profiling techniques predict inferred functional capacity without the need for costly, intensive full-scale metagenomics and without being limited to targeting select functional genes (Langille 2018).

Overall, validation is needed for phytoremediation of aromatic hydrocarbons in fractured rock. The objective of this study was to address temporally variable water use and contaminant uptake that affect quantification of phytoextraction rates, while also accounting for biodegradative processes with effective monitoring and performance evaluation techniques. In the present, year-long study (November 2016–November 2017), the efficacy of a small field-scale hybrid poplar phytoremediation system for remediation of an aromatic hydrocarbon (toluene) impacting a complex, fractured dual-porosity dolostone bedrock aquifer was evaluated. Emerging *in planta* contaminant passive sampling techniques and advanced, high-throughput metagenomic techniques were employed to assess phytoextraction performance and to elucidate root-associated biodegradation activity.

3.3 Materials and Methods

3.3.1 Site description and infrastructure

The investigated site was a manufacturing facility in a mixed residential/industrial neighborhood in urban South-Western Ontario, Canada that was operational between 1952 and 1990. Legacy subsurface toluene impacts were discovered in 1989 during decommissioning of outdoor storage tanks and buried product distribution lines that supplied toluene for use in the facility. Detailed site history and descriptions of characterization efforts to-date can be found elsewhere (Fernandes 2017; Roebuck 2018; BenIsrael et al. 2019; Wanner et al. 2019).

Surficial overburden primarily composed of sandy, cobbled soil is 2 m thick and overlays a Silurian-era fractured dolostone aquifer. Site hydrogeology and contaminant phase and mass distribution were previously characterized using the high-resolution Discrete Fracture Network (DFN) approach (Fernandes 2017). Three hydrogeological units (HGU1–3) were delineated based on vertical hydraulic head profiles. The shallowest unit, spanning 2–6.5 m below ground surface (bgs), has relatively high horizontal and low vertical hydraulic conductivity. Groundwater flows in a northern direction with velocities ranging from <0.01 to 3.0 m/day. The water table fluctuates seasonally between 2.1 and 3.0 m bgs, coincident with the bedrock-overburden contact, and is generally higher in spring compared to late summer and fall (Wanner et al. 2019). The primary source zone of toluene in groundwater occurs as residual LNAPL and in dissolved and sorbed phases within shallow bedrock in the north-eastern corner of the site (Figure 3.1A). A persistent groundwater plume is being sustained primarily through dissolution of residual LNAPL into groundwater. Approximately 95% of dissolved-phase toluene resides within the upper 2 m of the saturated zone. The bulk of toluene mass is located favorably in the saturated zone for plant-mediated uptake and biodegradation. A seasonally variable, spatially

homogeneous toluene vapor phase plume persists in the vadose zone as well, at concentrations significantly lower than in groundwater (BenIsrael et al. 2019).

In 2008, the site was planted with a 51-tree hybrid poplar (*Populus deltoides* × *nigra* OP-367) phytoremediation system (Figure B.1). Trees were planted overlaying the former tank and supply line area in nutrient-amended boreholes advanced to refusal up to approximately 2 m bgs at the bedrock-overburden contact, following established methodology (Ferro 2003). A network of six monitoring wells was installed in 2015 within the phytoremediation planted area, each with six monitoring ports up to a maximum depth of 10 m bgs (Figure 3.1). Three soil vapor probes were installed on site in 2016, each with three monitoring ports up to a maximum depth of 2.18 m bgs (Figure 3.1A). Average tree diameter at standard breast height was 15.97 cm in September 2016 and 17.57 cm in November 2017, before and after the present study period. Preliminary evaluation of groundwater utilization by the trees using stable water isotope analysis of ground and xylem water (Link et al. 2015) has shown evidence of uniform groundwater use by trees across the site (data not shown). Since installation of the phytoremediation system, toluene biodegradation in groundwater has been demonstrated at the site through compound-specific isotope analysis of toluene and detection of bacterial toluene degradation genes and transcripts (Roebuck 2018; Wanner et al. 2019). Both aerobic and anaerobic degradative processes are active in this aquifer, with the dominant processes shifting seasonally and with temporally changing groundwater redox conditions. Biodegradation has been demonstrated also in the vadose zone, with enhanced remedial activity occurring at depth (BenIsrael et al. 2019).

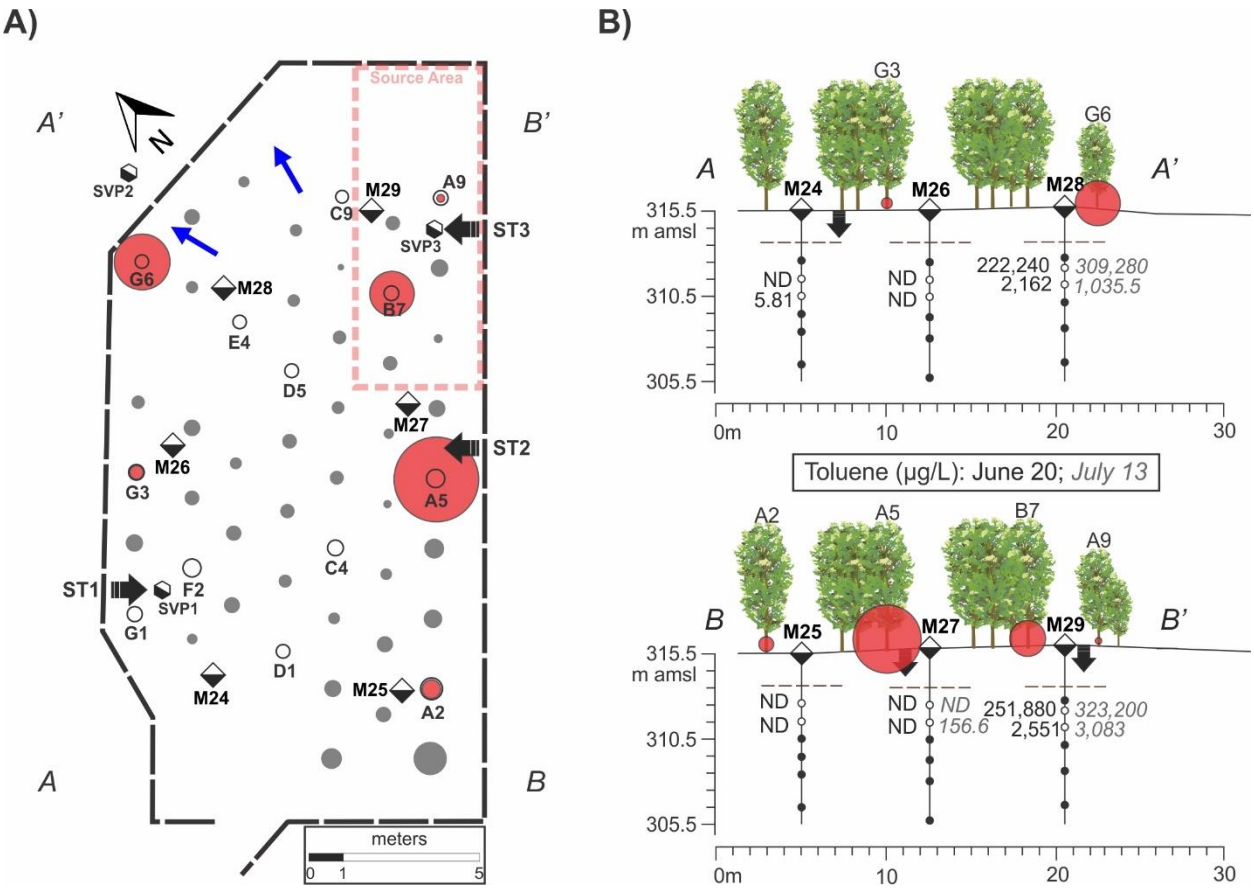


Figure 3.1: A) Phytoremediation area showing multilevel monitoring wells M24–M29 (half-shaded diamonds), multilevel soil vapor probes SVP1–3 (half-shaded hexagons), inferred direction of groundwater flow (blue arrows) in the upper hydrogeological unit (HGU1), toluene groundwater source area (red dashed box), unsampled trees (filled circles), trees instrumented with SPSs from June 20–July 11, 2017 (open circles), *in planta* toluene concentrations from SPS sampling (filled red circles), cross-sectional transects (*A–A'* and *B–B'*), and locations of November 2016 sampling trenches ST1–3 for root sampling (black arrows). Note: filled red circle sizes correspond to *in planta* toluene concentrations, with toluene detections ranging from 0.11 to 1.2 µg/L. B) Transects *A–A'* and *B–B'* cross-sections showing multilevel well sampling depths, depth to bedrock (dashed lines), November 2016 sampling trenches, toluene groundwater concentrations from June 20 (black text) and July 13 (gray, italicized text), 2017 sampling, and *in planta* toluene concentrations from SPSs (filled red circles) with corresponding tree IDs. ND = non-detect, vertical distances reported in m above mean sea level (m amsl).

3.3.2 Phytoextraction of toluene by hybrid poplars

3.3.2.1 Groundwater hydraulic head and toluene groundwater concentration measurements

Manual hydraulic head measurements, reported in m above mean sea level (m amsl), were taken with a coaxial water level meter in all water-bearing sampling ports of the six multilevel wells within the planted area on June 20 and July 13, 2017 (Figure 3.1). Groundwater samples for analysis of toluene were collected from the two uppermost water-bearing ports of all wells on June 20, 2017 and from select wells (M27, M28, and M29) on July 13, 2017. Toluene was quantified by GC-ECD (Agilent 7890A GC equipped with a DB-VRX column of 20 m, 0.18 mm ID, and 1 µm film thickness). The carrier gas was helium with a flow rate of 1 mL/min. The GC oven temperature for toluene analysis was held at 36 °C for 4 min, then ramped at 16 °C/min to 85 °C, further ramped at 30 °C/min to 210 °C and held for 3 min. This protocol had a method detection limit of 0.2 µg/L and a reporting limit of 1 µg/L.

3.3.2.2 Transpiration rates and weather

Poplar transpiration rates were quantified using Dynamax Inc. Thermal Dissipation Probes (TDPs) between May 20 and November 1, 2017, encompassing early leaf-out to pre-dormancy conditions. Probes were installed at 1.07 m above ground surface (ags) on south-facing sides of 5 trees (diameters ranging from 13.53 to 20.21 cm), equivalent to approximately 10% of the stand. Measurements were collected every 30 s and averaged over 15-minute intervals. Radial xylem symmetry of this poplar cultivar enabled installation of a single probe for each instrumented tree (Zalesny et al. 2006), and trees were deployed across the site to account for spatial heterogeneity of the stand (Kostner et al. 1998). Volumetric water use rates, V_T (L/hr), were calculated for each probe according to manufacturer instructions utilizing the Granier method (Granier 1985), and averaged. V_T per cm^2 basal area (transpirational cross-sectional area) rates were

calculated and basal areas, calculated using tree diameters at standard breast height measured on September 8, 2016, were used to estimate water use on a whole-stand and per-tree basis. Temperature was measured on site, while daily precipitation and hourly solar radiation data were obtained from regional weather stations (Climate ID 6144239; Elora Research Station 2017). Tree canopy photos were collected daily on site using a Moultrie Game Camera (Figure B.2).

3.3.2.3 Toluene in planta passive sampling and mass removal rates

In planta toluene concentrations were quantified following a 21-day equilibration period from June 20 to July 11, 2017 using polydimethylsiloxane solid phase samplers (SPSs) according to previously described methods (Shetty et al. 2014). Samplers were installed in 13 trees on south-facing sides in sealed trunk coreholes (4 cm x 0.5 cm) advanced at 0.5 m ags using a Haglöf Sweden 0.5 cm increment borer (Figure 3.1A). Tree diameters at height of instrumentation ranged from 15.7 to 26.7 cm (average of 19.1 cm), as measured on June 20, 2017. After the equilibration period, SPSs were collected into 20-mL vials with polytetrafluoroethylene-lined septa and stored on ice and in the dark until analysis. Field blanks were collected on site during SPS retrieval to evaluate background levels of toluene in air. Samplers were deployed concurrently in two trees (*Populus nigra* × *maximowiczii* NM-6, diameters of 20.05 and 19.1 cm) on a control plot with no known subsurface toluene impacts. The control plot is located 1.2 km from the study site and trees were planted there in 2005. Retrieved samplers were sent within 1 day to the Environmental Research Center at Missouri University of Science and Technology (Rolla, USA), where toluene concentrations were evaluated using solid phase microextraction gas chromatography (SPME-GC) (Limmer et al. 2011).

Toluene mass removal rates were calculated using the following equation:

$$M = V_{TB} \times C_t \times A_t$$

where M is mass removal rate (µg/day); V_{TB} is the volumetric water use rate per unit basal area (L/day• cm^2); C_t is the cumulative time-weighted average *in planta* concentration of detected toluene (µg/L); and A_t is the cumulative basal area of trees with detected toluene (cm^2).

3.3.3 Root DNA samples, amplicon library preparation, and sequencing

3.3.3.1 Root DNA samples

Poplar roots were collected in November 2016 from sampling trenches excavated using a CAT® 420 excavator with an 18-inch bucket at three locations (ST1–3) on the edges of the stand (Figure 3.1A), within a depth range of 0.97 to 1.93 m bgs, as described in BenIsrael et al. (2019). Briefly, 45 g of roots (*ca*. 1–5 mm diameter) on average were collected from excavated soil at each location. Roots were flash-frozen in the field using liquid nitrogen submersion for 3 min, then transported on dry ice and stored at −80 °C until further processing. Surface cleaning was performed by vortexing *ca*. 6 g of root material in 250 mL sterile 1 × phosphate-buffered saline (PBS), followed by shaking at 220 rpm at 10 °C for 20 min, then rinsing in 100 mL sterile 1 × PBS. Surface-cleaned roots (*ca*. 3 g) were homogenized using an MM400 Mixer Mill (Retsch GmbH) by agitating flash-frozen grinding jars twice at 25 Hz for 20 s. Total root DNA was extracted from root homogenate (0.3 g each in two pooled extractions) using a Mobio Laboratories Inc. PowerSoil® DNA Isolation Kit (BenIsrael et al. 2019).

Excavations at ST1 and ST2 were in the inferred upgradient direction from the source zone, in which toluene subsurface impacts were at low levels or not present on the basis of bedrock core, soil, and groundwater samples. The excavation at ST3 was within the source area, where toluene concentrations in bedrock core samples and shallow groundwater are the highest measured across the site (Fernandes 2017). As such, root samples collected from ST1 and ST2 (duplicate samples from each) were

designated as ‘low toluene (ST1/2) roots’ and those from ST3 (triplicate samples) as ‘high toluene (ST3) roots’.

3.3.3.2 Amplicon library preparation and sequencing

Amplicon library preparation for 16S *rRNA* genes in root tissue DNA samples was performed as described previously (Habtewold et al. 2018) and details are provided in Appendix B. Briefly, portions of the V4 region of bacterial 16S *rRNA* genes were amplified in duplicate samples using the modified 515F and 806R primer set (Table B.1) with Illumina adapter sequences (Caporaso et al. 2012; Apprill et al. 2015; Parada et al. 2016). A semi-nested PCR approach was used whereby duplicate samples were pooled and Illumina index tags were added in a second PCR, after which PCR products were purified, normalized to equimolar concentrations, and sequenced using Illumina MiSeq 2 × 250 bp paired end sequencing at the University of Guelph Advanced Analysis Centre, Genomic Facility (Guelph, ON).

3.3.3.3 Bioinformatic analysis

Raw FASTQ sequence files for bacterial 16S *rRNA* genes in root DNA samples were processed and analyzed using Qiime2 v. 2018.8 software (Bolyen et al. 2019). Forward and reverse reads of each sample were merged using the demultiplexing protocol for single-end reads barcoded according to the Earth Microbiome Project protocol (Caporaso et al. 2012). Sequence quality control was performed using the DADA2 pipeline for detecting and correcting sequencing errors (Callahan et al. 2016): a trim length of 35 (forward and reverse) and truncation lengths of 209 (forward) and 183 (reverse) were selected based on sequencing quality scores (sequencing quality control metrics presented in Table B.2). Sequences were grouped into sequence variants (SVs) based on 100% similarity. Sequence taxonomy was then classified using the Silva 132 classifier for 99% classification targeting the 515–806 region with seven-level taxonomy (Quast et al. 2013), and sequences were filtered to remove non-bacterial, mitochondrial,

and chloroplast sequences. Filtered sequences were rarefied to the smallest number of observed reads (29,370), then the mafft, mask, and FastTree protocols were used to generate rooted and unrooted phylogenetic trees of aligned representative sequences for use in diversity analyses. Sample diversity metrics were generated for α-diversity (Shannon Diversity Index, Pielou Evenness, Faith Phylogenetic Diversity, and number of observed sequence variants) and β-diversity (Bray-Curtis Dissimilarity). Finally, predictive evolutionary modeling of inferred community function in root DNA samples was performed using the PICRUSt2 software plugin in Qiime2 (Langille et al. 2013; Douglas 2019). The MetaCyc superpathway of aerobic toluene degradation (PWY-5183) functional feature (Caspi et al. 2014); Figure B.3) was targeted to assess predicted capacity for aerobic toluene degradation, and functional capacity is reported as the frequency of this functional feature relative to all predicted functional features.

3.3.3.4 Statistical analyses

Diversity profiles of poplar root DNA bacterial 16S *rRNA* sequences were generated for high toluene (ST3) and low toluene (ST1/2) roots and compared using Qiime2 software (Bolyen et al. 2019). α-diversity metrics were compared using the Kruskal-Wallis non-parametric one-way ranked analysis of variance (ANOVA) test. Multidimensional scaling of the generated β-diversity Bray-Curtis Dissimilarity matrix was performed using principal coordinates analysis (PCoA) to visualize main differences between samples in low-dimensional space. Differences between samples were evaluated using permutational multivariate analysis of variance (PERMANOVA) with pairwise comparison. Taxonomic and functional profiles of sequences were evaluated using STAMP v. 2.1.3 software (Parks et al. 2014). Differences in relative abundances over taxonomic levels were compared between high toluene (ST3) and low toluene (ST1/2) roots using Welch's t-test, and differences in relative predictive metabolic function were compared using a multiple group analysis of variance (ANOVA) test with Tukey-Kramer

pairwise comparison. Statistical significance in all cases was determined using a threshold of α=0.05.

3.3.3.5 Sequence accession

The unprocessed sequence reads of bacterial 16S *rRNA* genes obtained in this study have been deposited in NCBI's Sequence Read Archive (SRA) under project PRJNA551497 as FASTQ files with the accession numbers SRR9603253–SRR9603259.

3.4 Results and discussion

3.4.1 Poplar transpiration rates, weather, and groundwater hydraulic heads

Tree transpiration rates were evaluated, and air temperature, precipitation, and solar radiation were measured from May 20 to November 1, 2017 (Figure 3.2). Average volumetric water use (V_T) per cm^2 basal area ranged from 0.012 to 0.257 L/(day•cm^2) (average of 0.108 L/(day•cm^2)). Using these transpiration rates and tree basal areas from September 8, 2016, mean water use rates during the instrumentation period were estimated to be from 45.29 to 1,015.45 L/day (average of 425.40 L/day) for the stand and from 1.16 to 26.04 L/day (average of 10.91 L/day) for one tree of mean basal area. Rainfall and solar radiation patterns observed ranged from 0 to 28.6 mm/day (average of 1.85 mm/day) and 0 to 3.65 W/m^2 (average of 0.719 W/m^2), respectively. Air temperature was between 4.57 and 30.64 °C (average of 18.26 °C). Estimates of transpiration rates by the stand were consistent with prior reported values for trees of this cultivar and age, which typically range between 20.82 and 56.78 L/day (Ferro et al. 2001; ITRC 2009b). Fluctuations in the measured transpiration rates could be related to the local weather conditions. For example, the estimated whole-stand transpiration rate notably dropped from 468.58 L/day on July 12 to 65.15 L/day on July 13, 2017 (Figure 3.2). This corresponded to a decrease in peak solar irradiance from 1.57 to 0.72 W/m^2, a large rainfall event (28.6 mm) in the same period, and a drop in maximum measured air

temperature from 26.24 to 22.85 °C. Together, these results provide a reliable estimate of water use by the phytoremediation stand.

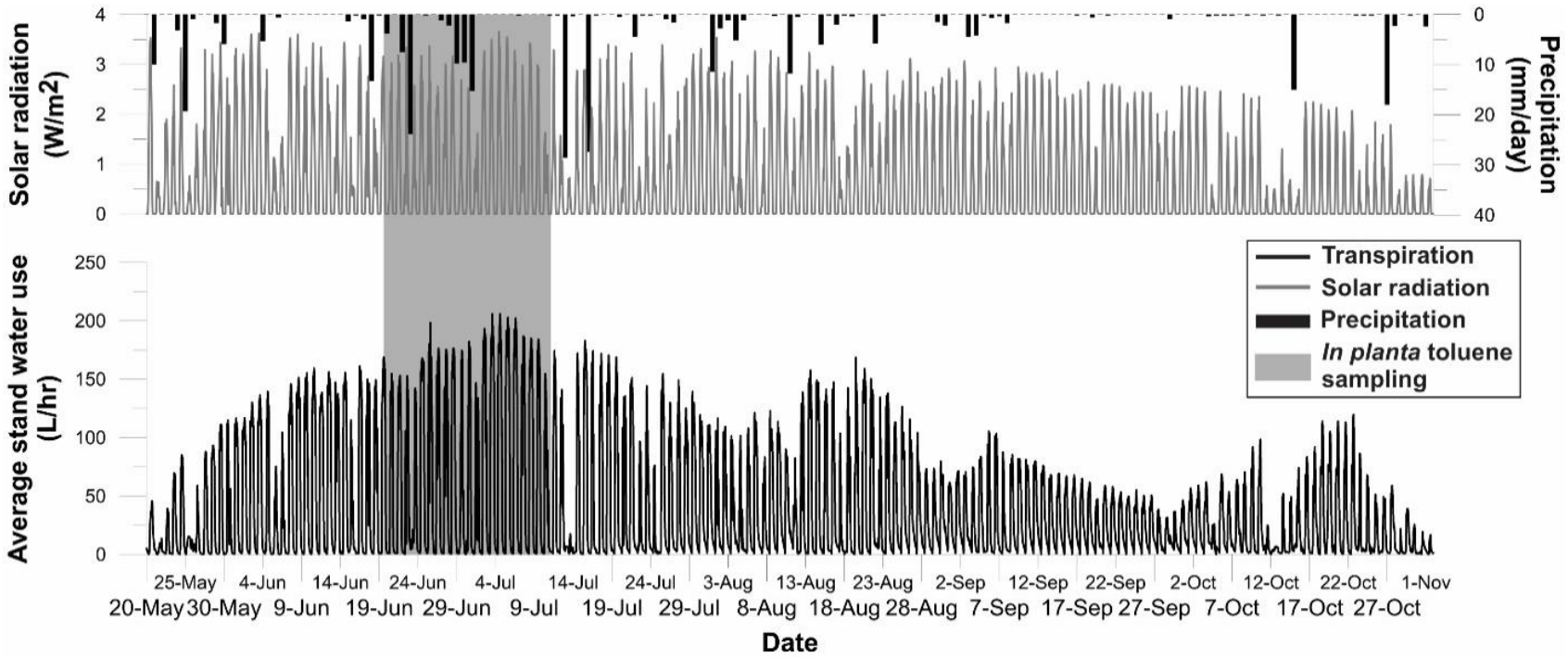


Figure 3.2: Transpiration, precipitation, and solar radiation assessed between May 20 and November 1, 2017. Transpiration was assessed on site and encompassed early leaf-out to pre-dormancy conditions for the stand. Daily precipitation and hourly solar radiation data were obtained from regional weather stations. *In planta* toluene sampling was conducted from June 20 to July 11, 2017.

Hydraulic head measurements were collected manually from the multilevel wells at the start and end of the *in planta* toluene monitoring period to assess water table height and site hydrology (Figure 3.3). In all wells, the highest hydraulic head values were measured in the uppermost 2–3 ports, within HGU1. This is most evident at M27, where the highest hydraulic head values within the planted area (313.38 and 313.37 m amsl in June and July, respectively) were measured in the uppermost port of that well. A slight decrease in water table height was observed between the June and July events, with uppermost hydraulic head values dropping by 0.01–0.06 m.

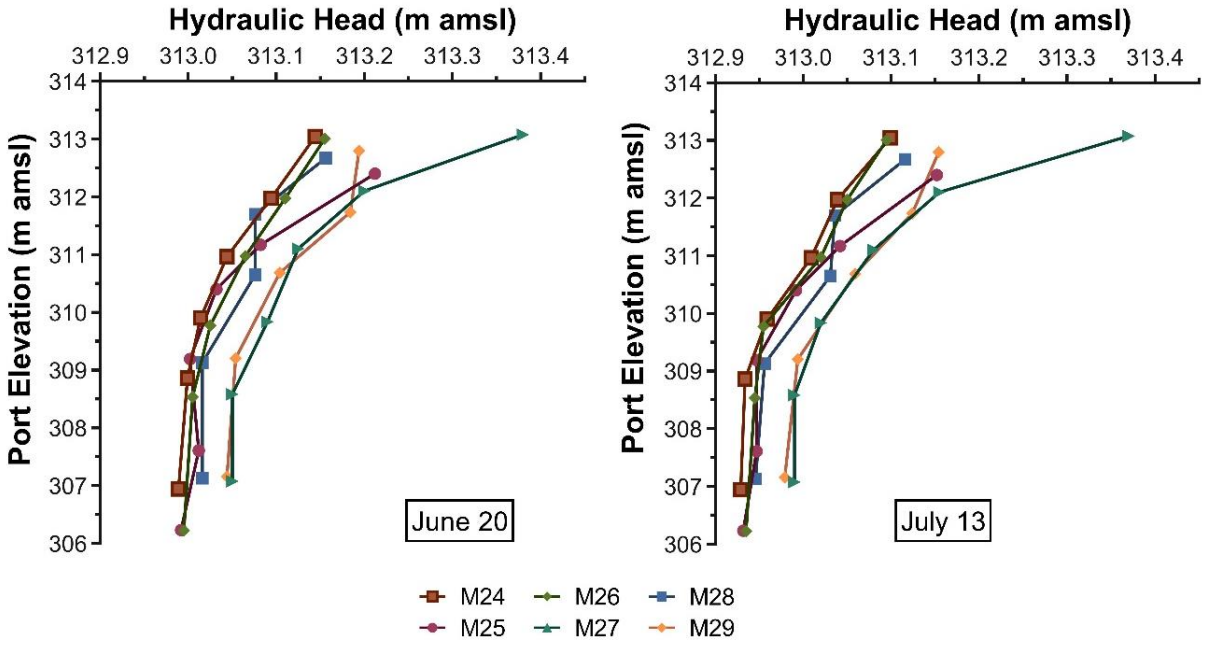


Figure 3.3: Manual groundwater hydraulic head measurements collected in all water-bearing sampling ports of multilevel wells on June 20 and July 13, 2017. Locations of multilevel monitoring wells (M24–M29) are provided in Figure 3.1A.

3.4.2 Toluene phytoextraction

3.4.2.1 Toluene concentrations in groundwater

Groundwater toluene concentrations were measured in the two uppermost water-bearing multilevel sampling ports on June 20, 2017 (all six wells) and July 13, 2017 (M27, M28, and M29) (Figure 3.1). Toluene concentrations detected were all below the maximum solubility limit of 521,000 µg/L (Sanemasa et al. 1982). Concentrations on June 20 ranged from non-detect (M25-27, M24 shallow port) to 251,880 µg/L (M29 shallow port). Maximal detected concentrations were greater on July 13, with the largest detection again at M29 (323,200 µg/L) and non-detects at M24-26 and the M27 shallow port. The greatest detections occurred within (M29) and downgradient (M28) of the source area and decreased by two orders of magnitude from the shallow to deeper sampling ports of M28 and M29. This is consistent with the source of the residual impacts being the product lines that were buried near the base of the overburden, and indicate that the poor vertical hydraulic conductivity conditions at the site result in strong attenuation of the plume in the vertical direction (Fernandes 2017; BenIsrael et al. 2019; Wanner et al. 2019). The measured toluene concentrations in groundwater provide an estimate of the detectable toluene mass that was available in groundwater for phytoextraction by the stand. Substantially greater mass was available to the trees in groundwater compared to the spatially ubiquitous vapor plume in the vadose zone, which had an average concentration, reported previously, of 0.011 µg/L_{vapor} on July 11, 2017 (BenIsrael et al. 2019).

3.4.2.2 In planta toluene concentrations and phytoextraction mass removal rates

Passive sampling techniques using SPSs were employed between June 20 and July 11, 2017, to establish time-weighted concentrations of toluene dissolved in xylem water. Toluene was detected in six of 13 instrumented trees, with concentrations observed between 0.11 and 1.2 µg/L, directly demonstrating active toluene uptake by the stand through phytoextraction (Figure 3.1, Table B.3). Field blanks collected on site

during sampler retrieval found no detectable background toluene, and toluene was not detected in off-site control trees. In poplar trees, core sampling direction typically correlates well to subsurface contaminant concentrations radially from sampled trees (Limmer et al. 2013). Trees were sampled uniformly on south-facing sides in this study, but it is possible that, if sampling was done in the known direction of highest groundwater impact (north to north-eastern), observed *in planta* concentrations might have been greater.

These results compare to a range of reported measurements of *in planta* toluene observed in previous studies assessing phytoextraction of BTEX constituents by poplars. Limmer et al. (2018) also utilized SPS passive sampling techniques in hybrid poplars (*P. charkowiensis* × *incrassate*) and detected a maximal toluene concentration of 52 µg/L. Their detection was an order of magnitude above what was found in this study and might be attributable to their use of TreeWell® planting techniques, whose facilitation of direct, isolated root interaction with impacted media can enhance mass removal processes (Gatliff et al. 2016). Wilson et al. (2013) detected toluene concentrations in hybrid poplars (*P. deltoides* × *nigra,* DN-34) of approximately 1 µg/L through tissue sampling methods; while this was not a field-level study and used a greenhouse plant reactor set-up, it simulated low, fluctuating water table conditions relevant to the present site.

Using daily V_T per cm^2 basal area for the stand (Section 3.4.1), the sum of detected *in planta* toluene concentrations, and summed basal areas for the six trees with detectable toluene in xylem water (Table B.3), daily toluene mass removal amounts for actively phytoextracting trees over the study period were estimated to be from 313 to 743 µg/day (average of 546 µg/day) (Figure 3.4). A total of 12 mg of toluene was calculated to be removed by sampled trees with detected toluene over the study period. Assuming a constant *in planta* concentration and using estimated transpiration rates between May 20 and November 1 as an approximation for water use over the growing season (early

leaf-out to pre-dormancy), a mass removal rate of 4.9×10^{-5} kg/tree/year was estimated for phytoextracting trees in the stand in 2017.

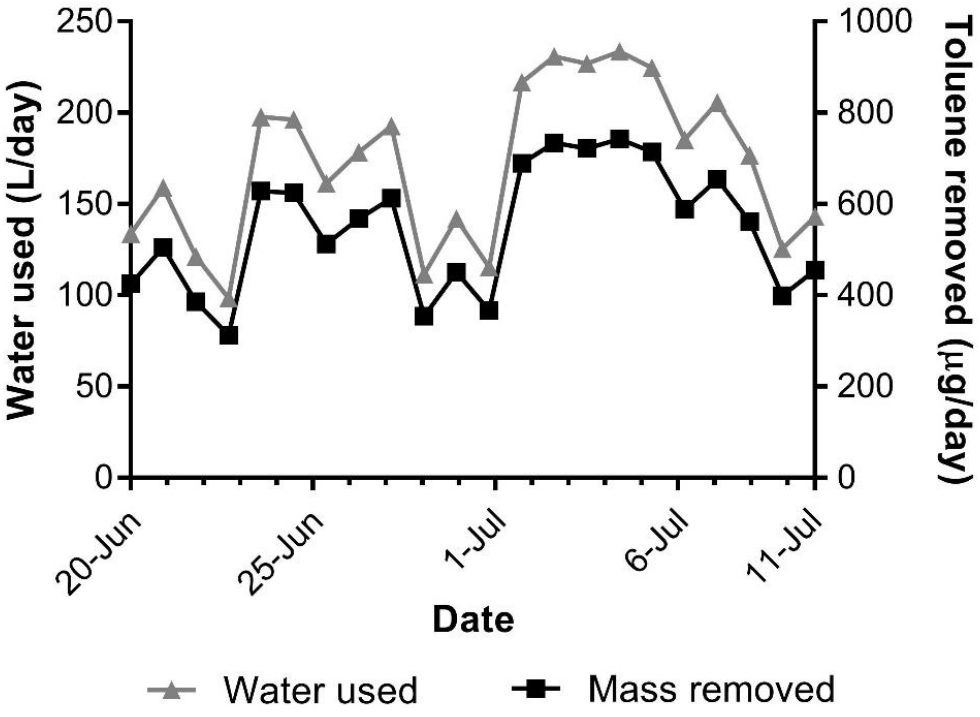


Figure 3.4: Cumulative daily water use and toluene mass removal rates calculated for six trees with detectable *in planta* toluene over the SPS sampling period (June 20–July 11, 2017).

The phytoextraction detection patterns observed showed that spatially variable phytoextraction is occurring on site. Trees with the greatest *in planta* toluene concentrations occurred above (G6 and B7) and ~upgradient/adjacent (A5) to the toluene groundwater source area (Figure 3.1). Tree A5 had the highest detected concentration (1.2 µg/L), despite growing above shallow groundwater with no detected toluene, while trees such as A9 and C9, overlying groundwater with a toluene concentration of 251,880 µg/L, had 0.11 µg/L and no detected toluene, respectively. Similarly, D5 and E4 had no detected toluene, despite proximity to groundwater with elevated toluene concentration at M28. Toluene was present in vapor phase within the vadose zone – 0.011 µg/L_{vapor} on

average on July 11, 2017 (BenIsrael et al. 2019) – and was likely a contributing source for toluene phytoextraction (Wilson et al. 2017). The greatest *in planta* concentrations coincided with the highest observed water table level, at M27 (Figure 3.3); however, toluene was not detected in groundwater from the shallowest port at this location. On a general level, the toluene concentrations observed *in planta* indicated higher toluene subsurface concentrations in the northern half of the planted area; though *in planta* detections were not highest in the source area (Figure 3.1A). Other studies have correlated *in planta* BTEX concentrations to concentrations in the subsurface (Schumacher et al. 2004; Limmer et al. 2011). However, variable factors influencing *in planta* toluene concentrations, such as fluctuating water table (Weishaar et al. 2009), hydraulic interactions with the trees (Matthews et al. 2003), dependence of tree transpiration on weather and diel cycles (Lundblad and Lindroth 2002), dissolution or volatilization of toluene from the sorbed mass (ITRC 2009a), and biodegradation occurring in each medium (ITRC 2009b) preclude correlations of this sort. Together, the spatial variability of the present phytoextraction results are likely best explained as a combination of hydraulic processes, coupled with variable biodegradation activity (and/or capacity) by plant-associated microbes (El-Gendy et al. 2009).

The difference between measured *in planta* toluene concentrations and predictable uptake concentrations for high toluene concentrations available in groundwater at this site directly support that biodegradation is occurring in the subsurface (Figure 3.1) (Burken and Schnoor 1998). This is consistent with Limmer et al. (2018), who identified rhizospheric degradation as a dominant end-fate for subsurface toluene, resulting in only a fraction of mass removal through phytoextraction. Active toluene biodegradation in rhizosphere soil has previously been shown at the present site (BenIsrael et al. 2019). Therefore, potential plant-associated biodegradation by root-colonizing bacteria occurring on site, and thereby influencing phytoextraction profiles, was explored further.

3.4.3 Root tissue microbial communities

To evaluate phyto-induced toluene biodegradation in relation to phytoextraction bacterial 16S *rRNA* genes in poplar root samples were sequenced to assess community structure and to predict bacterial biodegradation capacity. Samples were previously obtained at depth in November 2016 from three sampling trenches on site, representing high (ST3) and low (ST1/2) toluene conditions on the basis of the understanding of toluene distribution in the subsurface (Section 3.3.3.1; Figure 3.1). Diversity indices calculated for bacterial 16S *rRNA* amplicon sequences showed no significant differences in bacterial α-diversity between sampling locations or between high and low toluene conditions (Table 3.1). The bacterial β-diversity patterns observed in the present study implicate biodegradation of toluene as a strong, uniform selective pressure where the roots extend into the toluene-impacted subsurface, acting in tandem with other plant-specific plant-microbe interactions. Principal coordinates analysis (PCoA) of Bray-Curtis Dissimilarity metrics (59.16% variation in two dimensions) showed discrete clustering of samples on the basis of high toluene (ST3) and low toluene (ST1/2) designation (Figure 3.5). This indicated that distinct bacterial assemblages were established in the root microbiomes of trees overlying the source zone. A tighter grouping was observed between replicate high toluene (ST3) root samples, compared to the disparate grouping of low toluene (ST1/2) roots. Under unimpacted field conditions, poplar root endophytic communities typically form as unique assemblages compared to corresponding rhizosphere soil, rather than simply being opportunistic subsets recruited from dominant communities therein (Gottel et al. 2011). These distinct assemblages vary between poplar genotypes (Ulrich et al. 2008), and reflect the role of specific plant-microbe interactions in poplar trees for recruiting microbes capable of colonizing internal tissues (Gaiero et al. 2013). In the present study, low toluene (ST1/2) roots exhibited greater variability in their assemblages and were consistent with Shakya et al. (2013), who found that endophyte 16S *rRNA* sequences from poplar roots under unimpacted conditions typically exhibit significant variability between samples. Upon exposure to BTEX contaminants, degrader

organisms are enriched in hybrid poplar roots, and represent non-specific populations in plants (Moore et al. 2006; Barac et al. 2009). Overall, the present sequencing results illustrate substantial spatial variability of poplar root bacterial composition and marked uniformity of contaminant-exposed communities, and support conclusions of spatially-variable biodegradation occurring at the site.

Table 3.1: Diversity indices (±SD) of bacterial 16S *rRNA* gene sequences in roots from three sampling trench locations, categorized based on high toluene and low toluene designations.

	Sampling trench	
Index[a]	**Low toluene (ST1/2)**	**High toluene (ST3)**
Shannon Index	8.19±0.42	7.87±0.26
Pielou Evenness	0.84±0.01	0.81±0.02
Faith Phylogenetic Diversity	63.98±14.76	61.23±5.91
Observed Sequence Variants	857.25±249.91	834.00±39.95

[a] No significant differences found ($p>0.05$) between sampling trenches and high toluene (ST3) vs. low toluene (ST1/2) roots as assessed using the Kruskal-Wallis non-parametric one-way ranked analysis of variance (ANOVA) test.

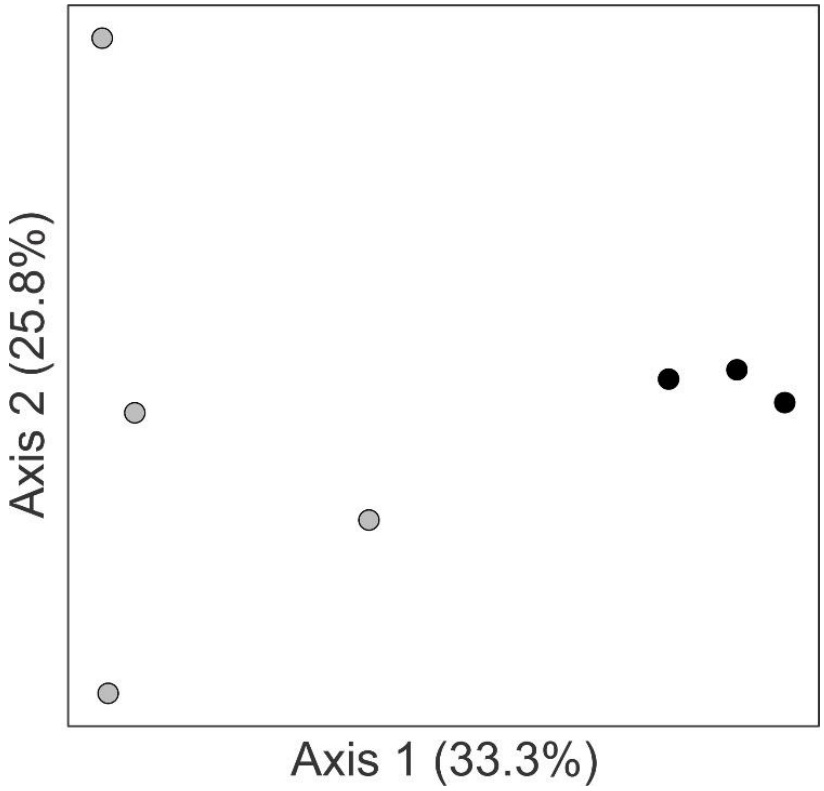


Figure 3.5: Principal coordinates analysis (PCoA) of Bray-Curtis Dissimilarity β-diversity metrics for bacterial 16S *rRNA* gene sequences in high toluene (ST3) vs. low toluene (ST1/2) roots. High toluene (ST3) roots (black, $n = 3$) differ significantly ($p<0.05$) from low toluene (ST1/2) roots (gray, $n = 4$) based on pairwise PERMANOVA comparison.

To elucidate which organisms were driving the observed bacterial compositional variability between high and low toluene conditions, differential sequence abundances were evaluated over numerous taxonomic levels between high toluene (ST3) and low toluene (ST1/2) roots (Figure B.4). On the phylum level, samples were composed mostly of Proteobacteria (41.7–47.7%), Actinobacteria (25.9–41.3%), and Bacteroidetes (7.3–13.7%), while on the class level Actinobacteria (22.4–37.9%), Alphaproteobacteria (19.6–28.4%), Gammaproteobacteria (11.4–18.1%), and Bacteroidia (7.3–13.7%) were dominant. Shakya et al. (2013) previously observed that Proteobacteria and Actinobacteria (62.4% and 23.9% of sequences, respectively) become enriched in

endosphere communities compared to corresponding rhizosphere soil, highlighting the substantial role of these taxa as poplar endophytes. These results were consistent with Ulrich et al. (2008), who found Proteobacteria and Actinobacteria represented 82% and 8.5% of sequences, respectively, in field-grown poplar roots. The present study supports these findings that Actinobacteria and Proteobacteria are dominant colonizers of poplar root microbiomes.

Root endophyte communities in poplars growing under BTEX-impacted field conditions were found previously to be dominated by Gammaproteobacteria (59% of isolates), particularly from the *Pseudomonas* genus (Moore et al. 2006). In the present study, however, predominance of Gammaproteobacteria was not observed in high toluene (ST3) compared to low toluene (ST1/2) roots. Ulrich et al. (2008) have proposed that the Gammaproteobacteria enrichment observed by Moore et al. (2006) reflected native degrader populations present in soil. Bacterial degraders in aromatic hydrocarbon-contaminated soils tend to shift to Actinobacteria dominance when organic matter and nutrient content are low, and to Proteobacteria dominance when high (Bell et al. 2013). Historical backfilling operations have occurred at the present site over the past 30 years, thus soil there might have lower organic matter and nutrient content than found at other poplar phytoremediation applications and therefore favored selection of Actinobacteria. Alternatively, Moore et al. (2006) utilized culture-based techniques to quantify endophyte organisms; inherent sampling bias of these techniques is well-established (Hirsch et al. 2010), and might have caused overrepresentation of Gammaproteobacteria in their study.

Together, these results support that poplar root endophyte communities are a combination of plant-specific and non-specific communities. Organisms are recruited into the root microbiome based on specific plant-microbe interactions and selected, in part, based on non-specific factors (i.e. catabolic genotype and biodegradation activity). In this study, differences were not necessarily apparent on higher taxonomic levels between

high (ST3) and low (ST1/2) toluene roots; therefore, organism abundances were evaluated on lower taxonomic levels.

Differences in sequence abundances between high toluene (ST3) and low toluene (ST1/2) roots were explored on the genus level to elucidate which organisms, potentially degraders, were enriched under high toluene conditions. Organisms classified to the *Pseudomonas* genus made up 0.23%–2.27% of sequences in high toluene (ST3) and 0.14%–0.85% in low toluene (ST1/2) roots. The greater relative proportion of *Pseudomonas* in high toluene (ST3) roots, although not statistically significant, is consistent with the demonstrated capacity of this genus for aromatic hydrocarbon degradation. Specifically, these organisms can degrade aromatic hydrocarbon compounds through side-chain oxygenation and ring mono- and di-oxygenation pathways, as well as degradation pathways for many alkanes, polycyclic aromatic hydrocarbons, and nitro- and chloro-aromatics (Palleroni et al. 2010). BTEX-derived enrichment of Gammaproteobacteria in poplar root endophyte communities has been attributed to this specialized capacity of *Pseudomonas* (Moore et al. 2006).

Six genera (Table B.4) with significantly greater relative abundances in high toluene (ST3) roots were identified in this study (Figure 3.6). Most notably, sequences classified to the *Streptomyces* genus represented 9.63%±1.06% of sequences, corresponding to an increase on average of 2.99% compared to low toluene (ST1/2) roots (6.64%±1.52%). *Streptomyces* contains plant growth-promoting species that affect nutrient uptake, plant growth, pathogen protection, and contaminant phytotoxicity in various plant species (Vurukonda et al. 2018). The genus was recently identified as a dominant colonizer of poplar roots (*P*. *trichocarpa* × *maximowiczii*) at a heavy metal-impacted field site, where substantial enrichment likely arose from prevalent, strong heavy metal resistance found within this genus (Durand et al. 2018). The role of *Streptomyces* in phytoremediation of aromatic hydrocarbons is relatively underexplored compared to heavy metal contaminants. Toluene degraders have been identified in the

Streptomyces genus (Balachandran et al. 2012; Mohamed et al. 2016), and Baoune et al. (2018) recently isolated 17 *Streptomyces* strains in root material obtained from three arid plant species growing in petroleum-impacted soil, with some isolates specifically capable of toluene biodegradation. Future work isolating *Streptomyces* spp. from the present site would be useful to demonstrate active toluene catabolism and support their involvement in toluene utilization *in situ*.

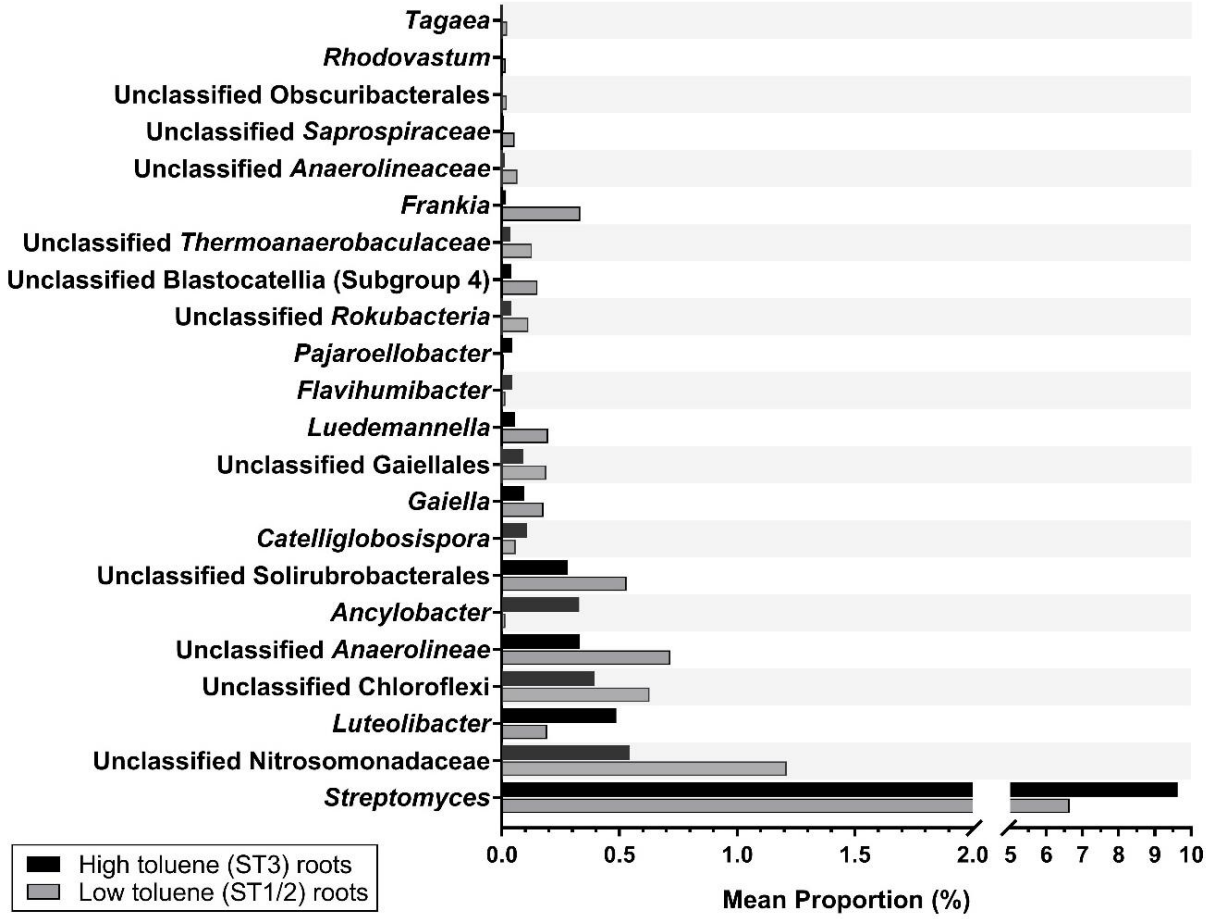


Figure 3.6: Bacterial 16S *rRNA* sequence relative abundance differences on the genus level between high toluene (ST3) and low toluene (ST1/2) roots. Only organisms with significant (Welch's two-sided t-test, $p<0.05$) differences in mean proportion between high toluene and low toluene roots are presented.

The significant enrichment of *Streptomyces* in high toluene (ST3) roots together with the relative enrichment of *Pseudomonas* support the potential for enhanced toluene biodegradation in roots overlying the source zone at the study site. Moreover, they suggest a substantial role for *Streptomyces* in poplar phytoremediation systems to address aromatic hydrocarbons, similar to that of *Pseudomonas* species at other sites. Indeed, Moore et al. (2006) noted their surprise at failing to find *Streptomyces* organisms among the root endophytes they isolated. Poplar root endophyte communities demonstrate dynamic seasonal variability (Shakya et al. 2013), while biodegradation rates (and therefore degrader competitiveness) are influenced by seasonal and diel variation in temperature, oxygen, and plant transpiration rates (Yadav et al. 2012; Yao et al. 2013; Landmeyer and Effinger 2016).

3.4.4 Taxonomic functional profiling

Differences in bacterial composition between high toluene (ST3) and low toluene (ST1/2) roots were found to occur at lower taxonomic levels (Figure 3.5; Figure 3.6). On the genus level, however, a substantial amount (36.32%) of sequence variants in root samples could not be classified. To overcome this gap in taxonomic classification, PICRUSt2 evolutionary modeling software was used to predict the relative frequency of aerobic toluene biodegradation capacity through known degradation pathways (Figure S3) between high and low toluene conditions. This analysis is made possible by the sufficient link between phylogeny and function that enables inferences of functional capacity of a metagenomic sample to be made using 16S *rRNA* sequences alone (Langille et al. 2013).

High toluene (ST3) roots had a significantly greater relative frequency (0.029%±0.003%) than low toluene (ST1/2) roots (0.016%±0.003%) for the aerobic toluene degradation functional feature (Figure 3.7). Thus, bacteria in roots overlying the source zone had a significantly greater predicted capacity for aerobic toluene degradation. These modelling results are supported by results previously reported for the

study site (BenIsrael et al. 2019). In that study, using quantitative PCR, a statistically significant ($p<0.05$) enrichment of bacterial toluene dioxygenase (TOD) genes was found in high toluene (ST3) compared to low toluene (ST1/2) root samples. Together, the predictive modelling and direct, quantitative, gene-specific results provide evidence that toluene impacts in the subsurface provide a selective pressure on root community structure through biodegradation. Furthermore, the enhanced biodegradative capacity of root bacterial communities in the source area provides a means of explaining the spatial patterns of *in planta* toluene concentrations (phytoextraction) across the site.

In using the PICRUSt2 software, the reliance on reference databases is recognized, and these are known to be dominated by Proteobacteria, particularly *Burkholderia cepacia*, *Pseudomonas mendocina*, *Pseudomonas putida*, and *Ralstonia pickettii* (Caspi et al. 2014). The model also has limited accounting for laterally transferred genetic elements, such as plasmid-encoded toluene degradation pathways (Taghavi et al. 2005), and for previously undiscovered degraders (Langille et al. 2013; Garza and Dutilh 2015). Together, these factors are likely to lead to an underestimation of the functional capacity of the degrader communities at this study site, but allow for comparisons of relative capacity to be made.

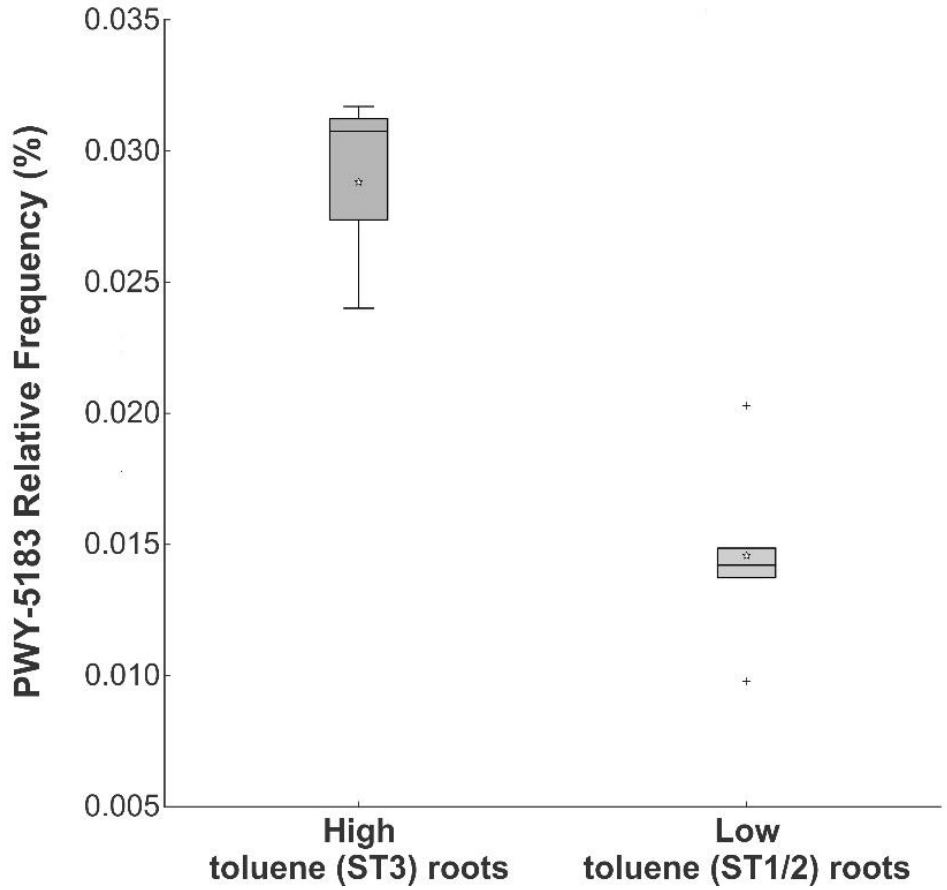


Figure 3.7: Relative frequency of the aerobic toluene degradation functional feature (MetaCyc PWY-5183) in root samples predicted using PICRUSt2 metagenomic functional modelling software. Functional feature relative frequency is calculated as the proportion of this feature relative to all predicted functional features. High toluene (ST3) roots have a significantly ($p<0.05$) greater relative frequency, as assessed using a multiple group analysis of variance (ANOVA) test with Tukey-Kramer pairwise comparison. Aerobic toluene degradation pathways targeted by this functional feature are defined in Figure B.3.

3.5 Conclusions

This study sought to quantify phytoextraction rates of toluene by hybrid poplars and to explore the relationship between root-associated bacterial communities and observed phytoextraction patterns. This is the first study to demonstrate quantitatively that hybrid

poplars extract an aromatic hydrocarbon (toluene) from a source zone in fractured rock. Toluene mass removal rates were quantified during peak seasonal activity by combining continuous transpiration measurement with *in planta* passive contaminant sampling. This combined approach was able to better account for temporal variation in transpiration and contaminant flux than is possible with instantaneous sampling methods.

The results indicated that the *in planta* concentrations of toluene found provided an approximation of subsurface impacts on a plot-wide scale, but not on a tree-by-tree basis. Thus, they demonstrate the potential of aromatic hydrocarbon passive sampling techniques to provide delineation of toluene in the subsurface – i.e. contaminant phytoscreening (Limmer et al. 2015) – on an approximate, qualitative basis. The results of this work highlight that on a tree-by-tree basis, the toluene concentrations measured in the transpiration stream at 0.5 m ags are affected by a number of processes. These include spatio-temporal variability in toluene concentrations in groundwater, temporally variable toluene concentrations in soil vapor in the vadose zone, water table depth, and the spatially-variable biodegradation capacity of plant-associated microbes.

The spatially-variable biodegradation capacity of the root-colonizing bacterial communities is a key finding of this work. A marked discrepancy was demonstrated between the low *in planta* toluene concentrations observed in the source area and the highly impacted groundwater there. This effect was explained by the corresponding enrichment of root-associated toluene degraders observed there. Root-colonizing bacterial communities were assessed, and their potential biodegradation activity predicted in this study, using metagenomic molecular analysis and functional evolutionary modeling of bacterial taxonomic marker genes in roots. A statistically higher proportion of sequences were from known microbial toluene degraders in roots collected from the source area than from roots collected where toluene impacts were low. Moreover, communities in the source area were more uniform and were distinct from low impact areas. This, along with previous results (BenIsrael et al. 2019) that quantitatively

demonstrated higher toluene degrading gene numbers in the high toluene area, illustrate how the presence of a contaminant provides a selective pressure on plant-associated populations, particularly at higher concentrations. Differences in root microbiome structure were not necessarily apparent on higher taxonomic levels between high and low contaminated conditions. This study contrasted previously observed selection of Proteobacterial poplar endophytes in BTEX phytoremediation systems and suggests that dominant phyla in contaminant-exposed root communities may depend on native soil physicochemical properties and microbial communities. Notable enrichment of *Streptomyces* was observed in roots collected under high toluene conditions, implicating this genus as a substantial poplar colonizer for BTEX phytoremediation.

Overall, this research provides evidence of the efficacy of hybrid poplar trees for active remediation of aromatic hydrocarbons at sites with complex hydrogeology and provides validation for emerging chemical sampling and metagenomic techniques that are essential for effective monitoring of contaminated sites and risk management programs. Future phytoextraction assessment approaches should, where possible, aim to also analyze biodegradation occurring along the transpiration stream to properly contextualize these parallel, overlapping processes in mass attenuation.

Acknowledgements:

This work was supported by BP, the University Consortium for Field-Focused Groundwater Contamination Research, and the Natural Sciences and Engineering Research Council of Canada (NSERC). Special thank-you to Alan Scheibner (BP Canada) for ongoing support and contributions to the project. Thank-you to Kamini Khosla, Jonathan Gaiero, Micaela Tosi, Jemaneh Habtewold, Dasiel Obregón Alvarez, Steve Chapman, Maria Gorecka, Patryk Quinn, James Hommerson, Brent Coleman, Kevin B. Ecott, Sean Jordan, Ian Renaud, and Steve Wilson for technical support.

4 Taxonomy and seasonality of tree stem phyllosphere microbiomes from hybrid poplars in a toluene phytoremediation system

This manuscript is in preparation

Keywords: BTEX aromatic hydrocarbons; endophyte; high-throughput sequencing; metagenomic molecular biological tools; phytoextraction; PICRUSt

4.1 Abstract

Hybrid poplars can actively remediate subsurface aromatic hydrocarbon contaminants, including the benzene, toluene, ethylbenzene, and xylene isomer (BTEX) family of compounds. Mass reduction processes in these systems are principally microbe-mediated and the contaminant profile strongly shapes community structure and function. While BTEX compounds are typically found in transpiration streams of poplars growing at impacted sites, little is known about the ecology and community structures of bacterial endophytes colonizing stem phyllospheres of poplars and how they might be influenced by aromatic hydrocarbon exposure. Culture-independent techniques were employed to characterize hybrid poplar stem endophyte communities at a field-scale toluene phytoremediation site in relation to measured toluene concentrations in groundwater, soil vapour, and quantified in tree transpiration streams. Generally, poplar stems were colonized predominantly by Gammaproteobacteria, Actinobacteria, Bacteroidia, Clostridia, Bacilli, Coriobacteriia, and Alphaproteobacteria. A site-wide seasonal increase in stem bacterial community diversity and richness was observed from early spring to late summer. This effect was greatest in a subset of trees growing in the more heavily impacted region of the site, which also had the highest measured *in planta* toluene

concentrations. An enrichment of *Rhodobacteraceae*, *Fibrobacter*, *Bacteroides ihuae*, *Rhizobiaceaee*, *Microbacteriaceae*, *Arcobacter butzleri*, *Pseudomonas*, *Pseudomonadaceae*, and *Elusimicrobium* phylotypes was observed in stem microbiomes of these trees compared to other trees in late summer. Moreover, these toluene-exposed stem microbiomes also had a greater predicted functional capacity for aerobic toluene catabolism, as assessed using taxonomic functional modeling.

4.2 Introduction

Hybrid poplar trees are commonly applied in phytoremediation of aromatic hydrocarbons, such as the benzene, toluene, ethylbenzene, and xylene isomer (BTEX) family of compounds (ITRC 2009b; Ferro et al. 2013; Landmeyer and Effinger 2016). Most mass reduction in these systems is believed to be microbe-mediated, occurring underground in the root-associated rhizosphere and root endosphere, wherein contaminants shape bacterial community structure (El-Gendy et al. 2009; Weishaar et al. 2009; BenIsrael et al. 2019). At sufficiently high source concentrations (and/or low biodegradation rates), BTEX compounds not attenuated in the rhizosphere can be taken up by poplar trees (i.e. phytoextracted), translocated along the transpiration stream, and can ultimately be phytovolatilized to the atmosphere through transpiration (Burken and Schnoor 1999; Limmer and Burken 2016; Limmer et al. 2018). Catabolic activity of endophytic bacteria colonizing phyllosphere (above-ground, internal) tissues along the transpiration stream can substantially reduce the amount of mass that is evapotranspired (Taghavi et al. 2005). While this can make the phytoremediation system more efficient, it can also reduce contaminant exposure risks, particularly when atmospheric release occurs in urban and residential areas.

Compared to contaminant-selected bacterial communities in rhizosphere and endosphere microbiomes (Jordahl et al. 1997; Moore et al. 2006; van der Lelie et al. 2009; Correa-Garcia et al. 2018), little is known about the ecology and community structures of

bacterial endophytes colonizing stem phyllospheres of poplars and how they might be influenced by aromatic hydrocarbon exposure. The stem phyllosphere of poplars is a distinct ecological niche compared to the root endosphere and other phyllosphere compartments (Beckers et al. 2017). As such, distinct bacterial communities exist there, and would be expected to habour a unique assembly of aromatic hydrocarbon-degrading organisms.

The present field study sought to explore the ecology of aromatic hydrocarbon-exposed endophytic bacterial communities in hybrid poplar stem phyllospheres under actively occurring phytoextraction conditions. Endophyte community structure was assessed using high-throughput sequencing of taxonomic marker genes in metagenomic stem tissue samples, collected in early spring and late summer from a mature toluene phytoremediation stand. Taxonomic profiling and functional prediction, as well as assessment of toluene phytoextraction and subsurface concentrations were employed to explore selection and enrichment of stem endophyte communities and contextualize their remediation activity.

4.3 Materials and Methods

4.3.1 Site description

This study was carried out at a manufacturing facility in a mixed residential/industrial neighborhood in urban South-Western Ontario, Canada. The site houses an ongoing phytoremediation program to address historic toluene subsurface impacts to a fractured bedrock aquifer system. Detailed site history and descriptions of characterization efforts to-date can be found in Section 2.3.1 and in Fernandes (2017), Roebuck (2018), and Wanner et al. (2019).

Briefly, toluene impacts were discovered in 1989 and believed to have come from buried product distribution lines connecting outdoor storage tanks to the facility. A network

of multi-level monitoring wells and multi-level soil vapour probes were installed within an approximately 220 m^2 fenced planted phytoremediation area on site in 2015 and 2016, respectively, to study contaminant mass and phase distribution. The wells each have six monitoring ports up to a maximum depth of 10 m below ground surface (bgs), while the soil vapour probes each have three monitoring ports up to a maximum depth of 2.18 m bgs. Wells and soil vapour probes used in this study are shown in Figure 4.1. Surficial overburden in the study area is composed primarily of sandy, cobbled soil, overlaying a Silurian-era fractured dolostone aquifer. Prior characterization has identified the primary source of toluene in groundwater as residual light, non-aqueous phase liquid (LNAPL) and in dissolved and sorbed phases within shallow bedrock in the north-eastern corner of the site (Figure 4.1). The water table fluctuates seasonally above and below the bedrock-overburden contact between 2.1 and 3.0 m bgs and is generally higher in spring compared to late summer and fall. A persistent groundwater plume is being sustained primarily through dissolution of residual LNAPL into groundwater. Approximately 95% of dissolved-phase toluene resides within the upper 2 m of the saturated zone.

A 51-tree hybrid poplar (*Populus deltoides* × *nigra* OP-367) phytoremediation stand was planted at the site in 2008 as a remediation and containment strategy. Trees were planted overlaying the former tank and supply line area in nutrient-amended boreholes advanced to refusal, following established methodology (Ferro 2003). Trees on site have been shown to be actively phytoextracting toluene from the subsurface, with highest concentrations phytoextracted at the north-eastern end of the site, above the source area (Chapter 3).

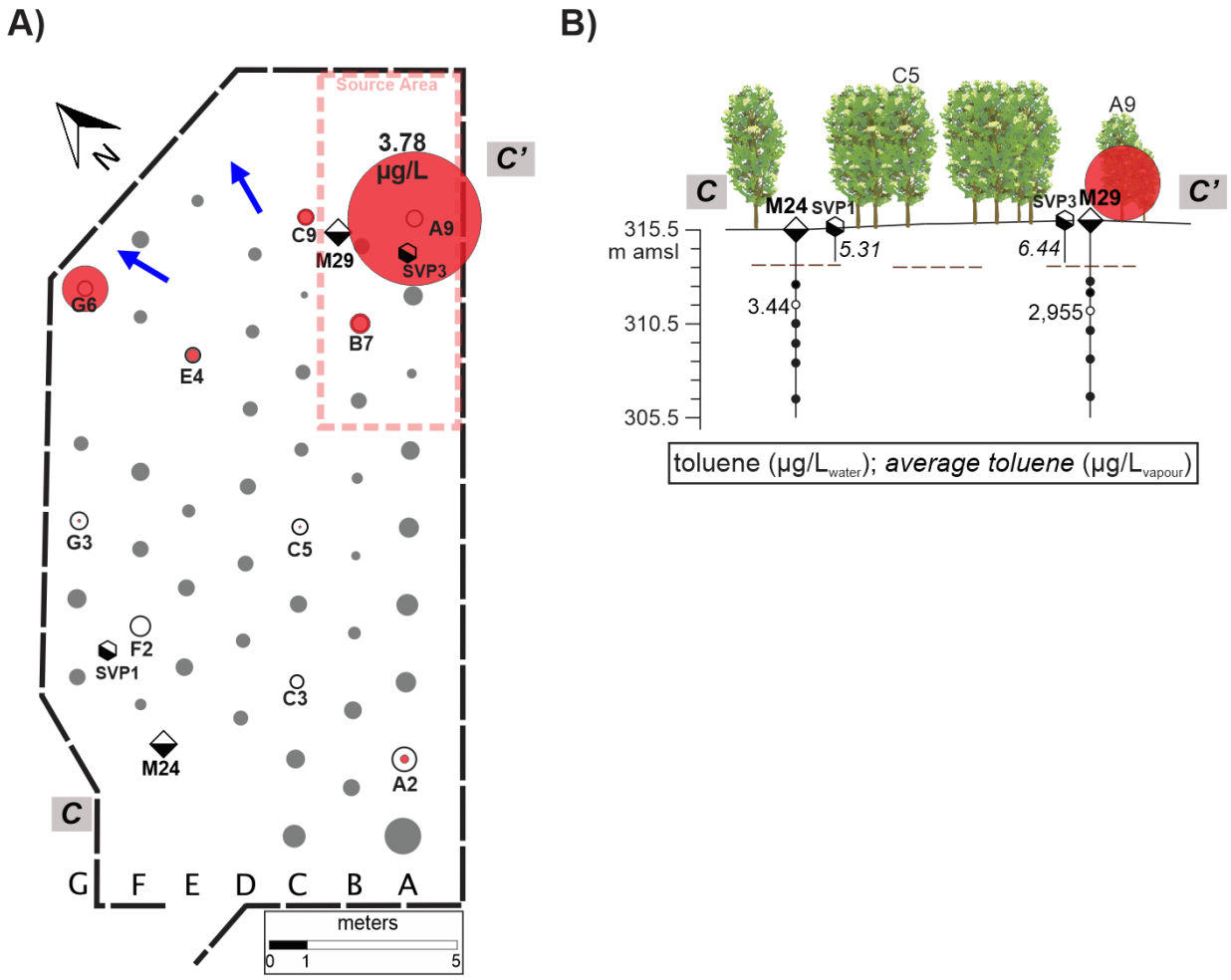


Figure 4.1: A) Phytoremediation site map showing fenced area, toluene source area, sampled (open circles) and unsampled (filled circles) poplar trees, monitoring wells M24 and M29 (half-shaded diamonds,), soil vapor probes SVP 1 and 3 (half-shaded hexagons), primary directions of shallow groundwater flow (blue arrows), and transect *C–C'*. Red shaded circles denote time-weighted concentrations of toluene in trees sampled August 27 to September 12, 2018 (non-detect–3.78 µg/L). B) Transect *C–C'* cross-section showing multilevel well sampling depths, depth to bedrock (dashed lines), August 27, 2018 toluene concentrations in groundwater (3.44 and 2,955 µg/L_{water}) and average concentrations in soil vapour (5.31 and 6.44 µg/L_{vapour}), and *in planta* toluene concentrations from trees A9 and C5 (filled red circles). Vertical distances reported in m above mean sea level (m amsl).

4.3.2 Groundwater and soil vapour sampling and toluene concentration analysis

Groundwater and soil vapour were measured and analyzed within the fenced planted area as described in BenIsrael et al. (2019) on August 27, 2018 from wells M24 and M29 and SVPs 1 and 3, respectively (Figure 4.1). Briefly, soil vapour was collected from all three sampling intervals into 1-L Tedlar® bags using an SKC XR500 air pump at a pump rate of 450 mL/min with dedicated 0.64 OD Teflon® tubing. Groundwater was collected from the uppermost water-bearing ports (3.97 and 4.32 m below ground surface at M24 and M29, respectively) using Geopump peristaltic pumps with dedicated 0.32 cm OD diameter Teflon® tubing. Toluene concentrations in groundwater and soil vapour were analyzed using a gas chromatograph coupled to a mass spectrometer (GC-MS).

4.3.3 Tissue collection and *in planta* toluene measurement

Stem tissue cores were sampled May 24 and August 27, 2018 to respectively assess baseline phyllosphere endophytic communities after trees reached full leaf-out following exit from winter dormancy (leaf budding observed May 8, 2018), and communities after summertime phytoextraction activity. Cores (0.5 x 4 cm) were collected using a Haglöf Sweden 0.5 cm increment borer from south-facing sides of 10 trees at 0.45 m above ground surface on May 24 and August 27, 2018 (Figure 4.1). Cores were collected into sterile Nasco Whirl-Pak™ bags on ice and preserved at -80 °C until processing. Tree diameters at height of sampling were between 14.32 and 28.33 cm (average of 20.41 cm).

Toluene concentrations at the sites of stem tissue samples were assessed according to methods by Shetty et al. (2014) and performed in BenIsrael et al. (2020). Briefly, polydimethylsiloxane solid phase samplers (SPSs) were installed for 16 days in sealed trunk coreholes between August 27 and September 12, 2018. Retrieved samplers were sent within 1 day to the Environmental Research Center at Missouri University of

Science and Technology (Rolla, USA), where toluene concentrations were evaluated using solid phase microextraction gas chromatography (SPME-GC) (Limmer et al. 2011).

4.3.4 Phyllosphere bacterial/archaeal gene sequencing and bioinformatics

Stem cores with outer bark removed were homogenized using an MM400 Mixer Mill (Retsch GmbH) by agitating flash-frozen grinding jars for 30 s at 25 Hz. DNA was extracted from on average 0.22 g of stem tissue homogenate (*ca.* 0.10 g on a dry weight basis) using a Dneasy PowerSoil® DNA Isolation Kit (Qiagen, Hilden, Germany) and gravimetric moisture content of tissue homogenate was determined. Extracted DNA purity was analyzed using a NanoDrop™ 8000 Spectrophotometer (Thermo Fisher Scientific Corp.) and quantities analyzed with a Qubit™ 4 Fluorometer (Thermo Fisher Scientific Corp.) using a Qubit™ dsDNA BR Assay Kit (ref. Q32850).

High-throughput DNA sequencing was performed targeting portions of the V4 region of bacterial/archaeal 16S *rRNA* genes. Amplicon preparation and sequencing were conducted at the McGill University and Génome Québec Innovation Centre (Montréal, QC, Canada) using the Illumina MiSeq PE300 amplicon sequencing platform. The 515F (5'-GTGYCAGCMGCCGCGGTAA) and 806R (5'-GGACTACNVGGGTWTCTAAT) primer set was used to amplify 16S *rRNA* genes as described in the Earth Microbiome Project protocol (Caporaso et al. 2012). Raw FASTQ sequence files were merged, processed, and analyzed using Qiime2 v. 2019.1 software (Bolyen et al. 2019). The DADA2 sequence quality control pipeline was employed using a trim length of 25 (forward and reverse) and truncation lengths of 220 (forward) and 196 (reverse) (Callahan et al. 2016). Taxonomic classification of sequence variants, grouped based on 100% similarity, was performed using the Silva 132 classifier for 99% classification targeting the 515–806 region with seven-level taxonomy (Quast et al. 2013). Sequences were filtered to remove non-bacterial/archaeal, mitochondrial, and chloroplast sequences by excluding based on taxonomic classification (10.1–94.5% of sequences removed), then rarefied to the smallest number of observed reads (3,598 reads). Rooted and unrooted phylogenetic

trees of aligned representative sequences were produced using the mafft, mask, and FastTree protocols. Sample diversity metrics were generated for α-diversity, including Shannon Diversity Index, Pielou Evenness, Faith Phylogenetic Diversity, and number of observed sequence variants, and for β-diversity (Bray-Curtis Dissimilarity).

Predictive evolutionary modeling of inferred bacterial community function in stem tissue samples was performed using PICRUSt2 analysis (Douglas et al. 2019) of 16S *rRNA* gene sequences. PICRUSt2 predicts potential function from 16S *rRNA* genes by (1) identifying nearest relatives with sequenced genomes; (2) predicting functional gene families in reference genomes likely to be conserved and therefore found in study organisms; (3) predicting the metabolic pathways present based on functional gene families predicted; and (4) calculating a frequency for each functional feature (pathway) based on sequence abundances in each sample. A Nearest Sequenced Taxon Index (NSTI) cut-off value of 2 was used for matching sequences to reference genomes. The MetaCyc pathways for aerobic toluene degradation via catechol (PWY-5178), *o*-cresol (PWY-5180), *p*-cresol (PWY-5181), and 4-methylcatechol (PWY-5182), intermediates (Caspi et al. 2014) were targeted to assess predicted capacity for aerobic toluene degradation. Functional capacity is reported as the frequency of these functional features relative to all predicted functional features.

The unprocessed sequence reads of bacterial/archaeal 16S *rRNA* genes obtained in this study have been deposited in NCBI's Sequence Read Archive (SRA) under project PRJNA551497 as FASTQ files with the accession numbers SRX7526665–SRX7526684.

4.3.5 Statistical Analyses

Diversity metrics were compared using Qiime2 v. 2019.1 software (Bolyen et al. 2019). α-diversity metrics were compared using the Kruskal-Wallis non-parametric one-way ranked analysis of variance (ANOVA) test. Multidimensional scaling of the generated β-diversity Bray-Curtis Dissimilarity matrix was performed using principal coordinates

analysis (PCoA) and differences between samples were evaluated using permutational multivariate analysis of variance (PERMANOVA) with pairwise comparison. Differences in relative abundances of taxa between sample groups were assessed in Graphpad Prism v. 7.04 using two-way analysis of variance (ANOVA) tests with post-hoc pairwise Holm-Sidak multiple comparisons tests. Differences between relative functional feature frequency between samples for predicted metagenomic function were evaluated using STAMP v. 2.1.3 software (Parks et al. 2014) using multiple group analysis of variance (ANOVA) tests with Tukey-Kramer pairwise comparisons. Statistical significance in all cases was determined using an α=0.05 threshold.

4.4 Results and discussion

4.4.1 Groundwater, soil vapour, and *in planta* toluene concentrations

Toluene concentrations in shallow groundwater and soil vapour were assessed on August 27, 2018 using two monitoring wells and two soil vapour probes (Figure 4.1) to measure phyto-available toluene prior to the *in planta* toluene monitoring period. Toluene concentrations in shallow groundwater were substantially higher (2,955 µg/L_{water} at M29) near the source area (north-eastern end) compared to the measured concentration at the south-western end of the site (3.44 µg/L_{water} at M24). Meanwhile, soil vapour toluene concentrations were by comparison relatively consistent across the site, ranging from 5.06 to 7.95 µg/L_{vapour} and averaging 5.87 µg/L_{vapour}. *In planta*, time-weighted toluene concentrations over the monitored period (August 27–September 12, 2018) were between non-detect and 3.78 µg/ L_{water} (greatest in A9), showing a similar trend to groundwater with the greater detection occurring at the north-eastern end of the site (Figure 4.1). These concentration profiles in all analyzed media were consistent with previous on-site patterns reported by BenIsrael et al. (2019), Wanner et al. (2019), and BenIsrael et al. (2020).

4.4.2 Stem phyllosphere endophytic community structure

Differences in α- and β-diversity were observed in the stem phyllosphere endophytic community between May and August samples marking a seasonal shift in microbiome community structure (Figure 4.2 and Table 4.1). Principal coordinates analysis (PCoA) of Bray-Curtis Dissimilarity metrics (40.92% variation in two dimensions) showed significant clustering of samples on the basis of sampling date (PERMANOVA, pseudo-F = 3.716, P = 0.001, Figure 4.2). Overlap in endophyte community relatedness was observed for some trees between these two time periods, while the significant differences in β-diversity between May and August samples were driven by a discrete cluster of August samples (A9, B7, C9**,** and G6, see locations in Figure 4.1). These samples, which were designated as the 'August 679-Cluster', had microbiomes that were significantly different from microbiomes in all other samples (PERMANOVA, pseudo-F = 4.964, P = 0.001, Figure 4.2). A significant increase in bacterial/archaeal diversity, richness, and evenness was observed from May to August, and the August 679-Cluster samples showed significantly greater α-diversity than all other samples (Table 4.1).

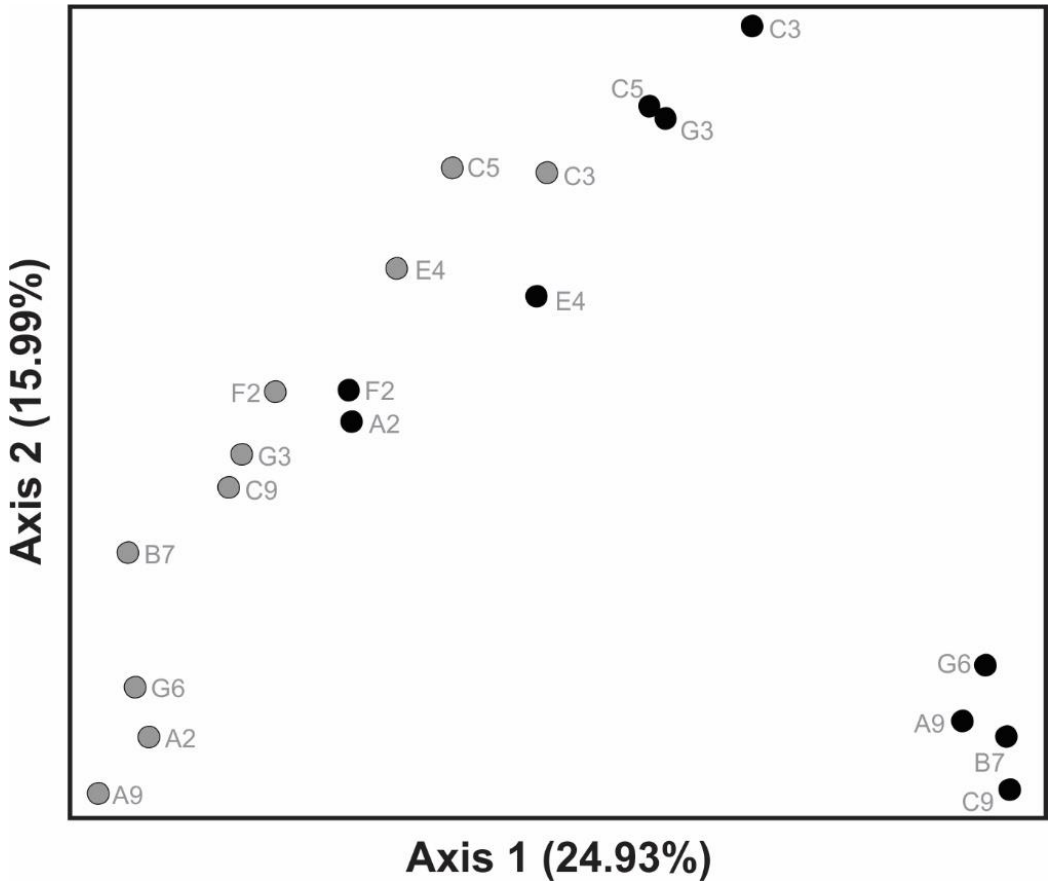


Figure 4.2: Principal coordinates analysis (PCoA) of Bray-Curtis Dissimilarity β-diversity metrics for bacterial/archaeal 16S *rRNA* gene sequences from May (grey, *n* = 10) and August (black, *n* = 10) 2018 stem tissue samples. Samples differ significantly (P<0.05) by sampling date based on pairwise PERMANOVA comparison. Four August 2018 samples in bottom right (from trees A9, B7, C9, and G6) are designated as the 'August 679-Cluster' and differ significantly (P<0.05) from all other samples. See Figure 4.1 for tree locations.

Table 4.1: Stem phyllosphere microbiome 16S *rRNA* gene α-diversity (Shannon Index, Pielou Evenness, Faith Phylogenetic Diversity, and Observed Sequence Variants).

Sampling date[a]	Tree	August 679-Cluster?[b]	Shannon Index	Pielou Evenness	Faith Phylogenetic Diversity	Observed Sequence Variants
May	A2	No	1.88	0.28	10.01	96
	A9	No	1.48	0.29	3.95	33
	B7	No	2.93	0.49	7.21	65
	C3	No	5.13	0.66	18.93	228
	C5	No	2.39	0.38	10.55	83
	C9	No	3.85	0.59	8.36	89
	E4	No	4.19	0.60	11.16	129
	F2	No	2.45	0.36	11.50	109
	G3	No	4.08	0.59	12.32	124
	G6	No	2.77	0.45	8.55	71
August	A2	No	4.69	0.70	10.53	107
	A9	Yes	6.59	0.90	18.75	157
	B7	Yes	6.78	0.86	22.32	231
	C3	No	5.02	0.69	13.34	150
	C5	No	5.16	0.72	12.51	148
	C9	Yes	6.82	0.84	23.13	287
	E4	No	4.72	0.71	9.14	99
	F2	No	3.85	0.58	10.67	98
	G3	No	5.28	0.72	15.28	163
	G6	Yes	6.80	0.86	24.68	239

[a] All α-diversity metrics are significantly ($P<0.05$) greater in August ($n = 10$) samples compared to May ($n = 10$).

[b] August 679-Cluster microbiomes defined according to distinct Bray-Curtis Dissimilarity grouping (Figure 4.2). All α-diversity metrics are significantly ($P<0.05$) greater in August 679-Cluster microbiomes ($n = 4$) compared to non-August 679-Cluster microbiomes ($n = 16$).

The apparent seasonal variability in endophyte community structure observed was not unexpected. Phyllosphere endophyte colonization in plants is influenced by many temporally-variable factors, such as temperature (Ju et al. 2006) and organic and inorganic compound composition of tree sap (Furukawa et al. 2011). Indeed, the endophyte microbiome is determined by environmental conditions and plant life stages (Kandel et al. 2017). This effect has also been observed in phytoremediation systems. At an urban petroleum hydrocarbon and chlorinated solvents-impacted site, Shen and

Fulthorpe (2015) found that stem endophyte community structure shifted substantially across seasons in maple and elm trees.

In August, similarities in the composition of the stem endophyte microbiomes from the August 679-Cluster trees implies that they might be influenced by a common selective pressure (i.e. toluene), reflecting the gradient of subsurface toluene impacts across the site. The August 679-Cluster trees are all found at the north-eastern end of the site, proximal to the bulk of dissolved phase toluene mass in groundwater and where *in planta* toluene concentrations were highest (Figure 4.1). Therefore, spatially-variable toluene phytoextraction and *in planta* exposure likely selected for distinct communities in August 679-Cluster trees. This is consistent with previous observation at this site of distinct root microbiomes in trees overlaying the source area compared to locations with lesser toluene availability (BenIsrael et al. 2019; BenIsrael et al. 2020).

Differences in species richness were also notable between August 679-Cluster and non-August 679-Cluster samples. Rare 16S *rRNA* gene sequences (i.e. in low relative abundance) made up a greater proportion of total sequences in the August 679-Cluster (38.54% ± 7.09%) compared to non-August 679-Cluster samples (8.67 ± 4.39). This is contrary to expectation, being that hydrocarbon exposure often produces a decrease in species richness and a shift in dominant phylotypes through selection and enrichment of degraders (Yang et al. 2016). Richness and diversity are typically lower in phyllosphere compartments compared to rhizosphere and root endosphere microbiomes (Beckers et al. 2017; Wang et al. 2019). The observed increase in richness perhaps speaks to a broad degradation capacity of stem endophytes in these trees either innately (Moore et al. 2006) or potentially through horizontal gene transfer (Taghavi et al. 2005). Alternatively, considering that the phyllosphere is more nutrient-limiting than the rhizosphere/root endosphere (van der Lelie et al. 2009), toluene might have supplied degraders with an alternative carbon source and offered non-degraders better access to plant-derived carbon sources.

4.4.3 Stem microbiome taxonomy

Apparent differences in poplar phyllosphere communities across the site were explored further on the taxonomic level. Poplar stem tissue samples on the class level comprised predominantly of Gammaproteobacteria, Actinobacteria, Bacteroidia, Clostridia, Bacilli, Coriobacteriia, and Alphaproteobacteria, and differed significantly between months (Figure 4.3A). Specifically, a decrease in average relative abundance of Gammaproteobacteria (66.07% to 22.54%) and increases in Bacteroidia (3.05% to 13.98%) and Clostridia (3.49% to 13.54%) occurred from May to August 2018. Proteobacteria and Actinobacteria are commonly observed as dominant taxa in poplar phyllospheres (Moore et al. 2006; Ulrich et al. 2008; Beckers et al. 2017), while Bacteroidetes and Firmicutes have also been observed in high abundance in poplar stems (Wang et al. 2019).

Dominant phylotypes in stem microbiomes, grouped based on common classification at the lowest resolved taxonomic level, showed a marked change from May to August and reflected a potential biotic stress occurring in these trees (Figure 4.3B). A large proportion of sequences found were classified to the *Dickeya* genus and the *Enterobacteriaceae* family in May (37.84% and 16.94%, respectively, on average), but these phylotypes were in lower average abundance in August (11.44% and 5.08%, respectively). The most abundant *Dickeya*-classified sequence (up to 77.36% abundance) had 100% homology to various *Dickeya* species, including *D. fangzhongdai, D*. *chrysanthemi*, *D*. *aquatica*, *D*. *dadantii* subsp. *dieffenbachiae*, *D*. *zeae*, *D*. *dianthicola*, and *D*. *paradisiaca*. The most abundant *Enterobacteriaceae* sequence (up to 55.97% abundance) was 98.35% similar to the same *Dickeya* species and to *Brenneria roseae* and *Brenneria nigrifluens*, while another sequence had 98.35% homology to *Lonsdalea populi* and other *Enterobacteriaceae* genera. *Lonsdalea populi* is the causative agent of poplar canker disease (Li and He 2019). *Brenneria roseae* and *Brenneria nigrifluens* are pathogens of oak and walnut trees, respectively, but are related to the poplar pathogen

Brenneria populi (Li et al. 2015). *Dickeya* spp. is considered a broad-host-range plant pathogen (Ma et al. 2007) causing various canker diseases (Alic et al. 2019); however, it has not previously been identified as a pathogen of poplar trees. Interestingly, endophytic *Dickeya* was found to increase in abundance in roots of willow trees (*Salix* spp.) exposed to petroleum hydrocarbons compared to non-exposed roots in a phytoremediation system (Tardif et al. 2016). Both *Populus* and *Salix* are members of the *Salicaceae* family and employed similarly in phytoremediation given their genetic and physiological resemblances (Rogers et al. 2019). It is unclear if presence of *Dickeya* spp. in stem microbiomes at the present site implies pathophysiology. Regardless, the presence of these phylotypes represents a potential pathogen presence in stem microbiomes of trees at this site, that apparently decreased in general from the early growing season to late summer.

A notable exception to this general trend occurred in tree E4, where at least 39.98% of sequences were classified as or were related to *Dickeya* spp. (97.53%–100% similarity). Tree E4 grew in the north-eastern end of the site; though it had measurable quantities of toluene in its transpiration stream in the August–September monitoring period (0.33 µg/L; Figure 4.1), its endophytic community in August did not group with the August 679-Cluster microbiomes (Figure 4.2). The higher relative proportion of suspected pathogenic bacteria in this tree likely explains why toluene exposure did not similarly shape endophyte community structure in this tree in August.

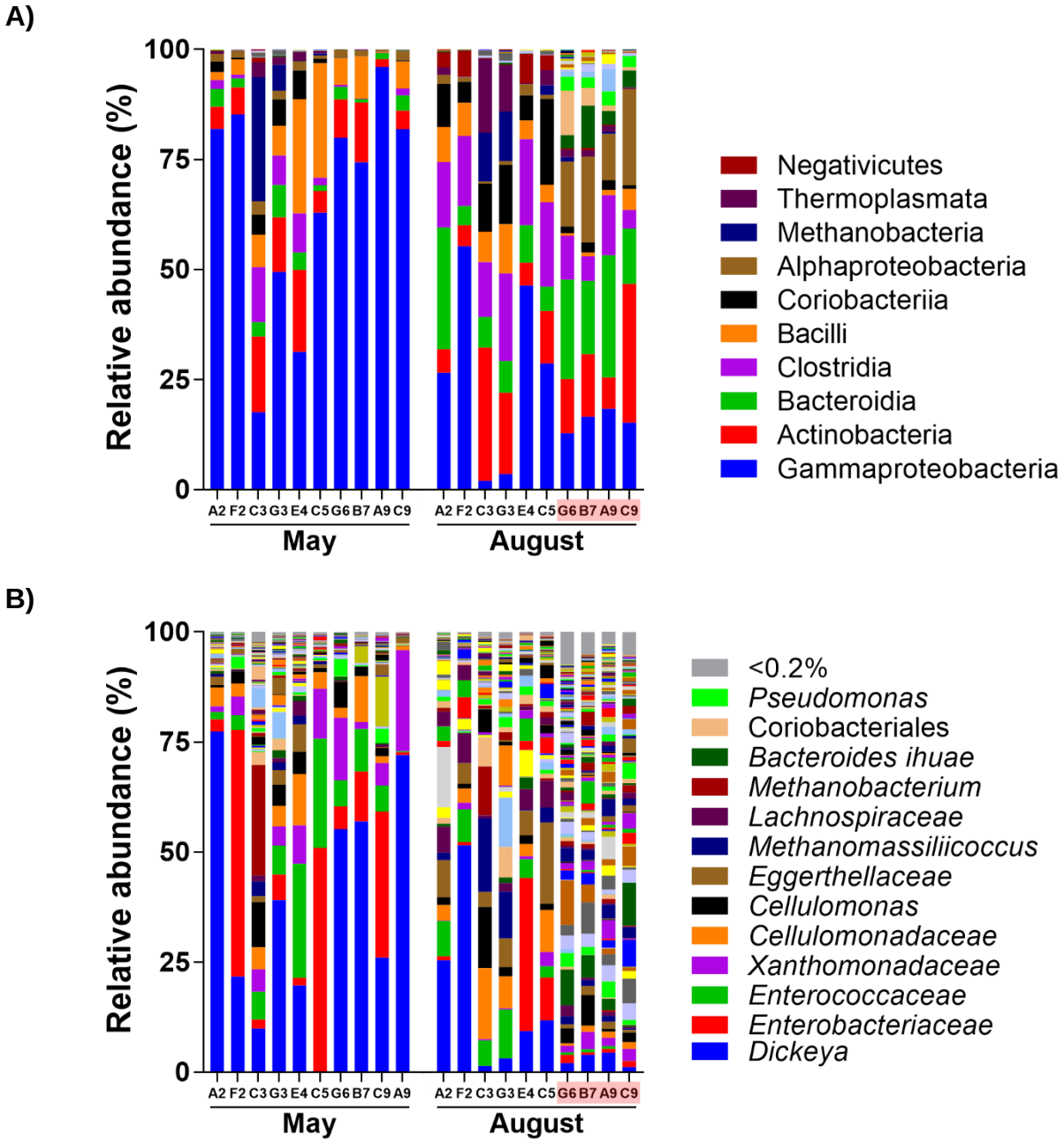


Figure 4.3: Relative abundance of 16S *rRNA* sequences in tree stem tissue samples grouped by class (A) and common classification to the lowest taxonomic level resolved (B). Taxa names displayed are lowest classification assigned to the genus level. The August 679-Cluster, as identified by β-diversity profiles (Figure 4.2), are indicated by shading of tree labels.

4.4.4 Toluene degrader enrichment

The effects of apparent toluene selection in the August 679-Cluster samples were explored further using functional taxonomic modeling. PICRUSt2 evolutionary modeling software was used to predict the relative frequency of aerobic toluene biodegradation capacity through known degradation pathways in metagenome samples. In August 2018, predicted relative capacity for aerobic toluene biodegradation through *p*-cresol (Figure 4.4A) and catechol (Figure 4.4B) intermediate pathways (Parales et al. 2008; Caspi et al. 2014) was significantly greater for August 679-Cluster microbiomes compared to other August samples. This observed enrichment of predicted toluene degraders in the August 679-Cluster samples supports that distinct endophyte community structure in stem tissue of these trees was a result of greater exposure to toluene.

Table 4.2: Differences in 16S *rRNA* gene relative abundance grouped by common classification between A9, B7, C9, and G6 (August 679-Cluster) and all other trees for August 2018 samples.

	August 679-Cluster		Non-August 679-Cluster		% difference in
Lowest classification	**mean (%)[a]**	**SD**	**mean (%)[b]**	**SD**	**means[c]**
Rhodobacteraceae	4.420	2.348	0.038	0.078	**4.382**
Fibrobacter	4.073	4.292	0.001	0.001	**4.073**
Bacteroides ihuae	4.020	3.294	0.753	1.047	**3.267**
Rhizobiaceae[d]	3.356	0.441	0.201	0.258	**3.154**
Microbacteriaceae	3.133	2.053	0.100	0.147	**3.033**
Arcobacter butzleri	2.772	0.419	0	0	**2.772**
Pseudomonas	2.402	1.207	0.201	0.359	**2.202**
Pseudomonadaceae	2.010	0.867	0.082	0.180	**1.928**
Elusimicrobium	1.948	2.193	0.006	0.015	**1.942**
Xanthomonadaceae	2.504	1.093	1.037	1.183	1.467
Flavobacterium	1.400	1.397	0.022	0.039	1.378
Tannerellaceae	1.925	1.673	0.621	0.696	1.304
Sphingobacteriaceae	1.161	1.056	0.154	0.191	1.007
Burkholderiaceae	1.011	0.797	0.034	0.057	0.977
Chthoniobacterales	0.931	0.425	0.024	0.049	0.907
Aeromicrobium	0.840	0.474	0	0	0.840
Rhizobiaceae	0.844	0.401	0.018	0.035	0.826

Paludibacter	0.798	0.431	0	0	0.798
Pseudoxanthomonas	0.795	0.329	0.017	0.043	0.778
Paludibacteraceae	1.205	1.277	0.444	0.577	0.761
Micrococcales	0.916	0.542	0.185	0.247	0.731
Fluviicola	0.725	0.609	0.016	0.038	0.709
Dysgonomonas	1.010	0.959	0.472	0.885	0.538
Christensenellacea	0.896	0.480	0.445	0.379	0.451
Pajaroellobacter	0.401	0.225	0	0	0.401
Cerasicoccus	0.397	0.255	0	0	0.397
Taibaiella	0.392	0.391	0.005	0.012	0.387
Rickettsiales	0.381	0.286	0.001	0.003	0.380
Clostridiales	0.455	0.333	0.181	0.101	0.274
Acinetobacter	0.280	0.130	0.011	0.013	0.268
Anaerocella	0.559	0.366	0.292	0.287	0.267
Paraburkholderia tropica	0.212	0.114	0.005	0.012	0.207
Terrimicrobium	0.194	0.047	0	0	0.194
Taibaiella	0.184	0.128	0.007	0.017	0.178
Dysgonomonas	0.229	0.295	0.061	0.149	0.168
Opitutaceae	0.164	0.046	0	0	0.164
Clostridiales	0.252	0.177	0.087	0.110	0.164
Chryseobacterium	0.162	0.031	0.005	0.012	0.157
Sandaracinobacte	0.141	0.061	0.007	0.015	0.134
Ruminococcaceae	0.374	0.265	0.246	0.242	0.128
Christensenellaceae	0.149	0.181	0.023	0.037	0.126
Ruminococcaceae	0.112	0.073	0	0	0.112

[a] $n = 4$

[b] $n = 6$

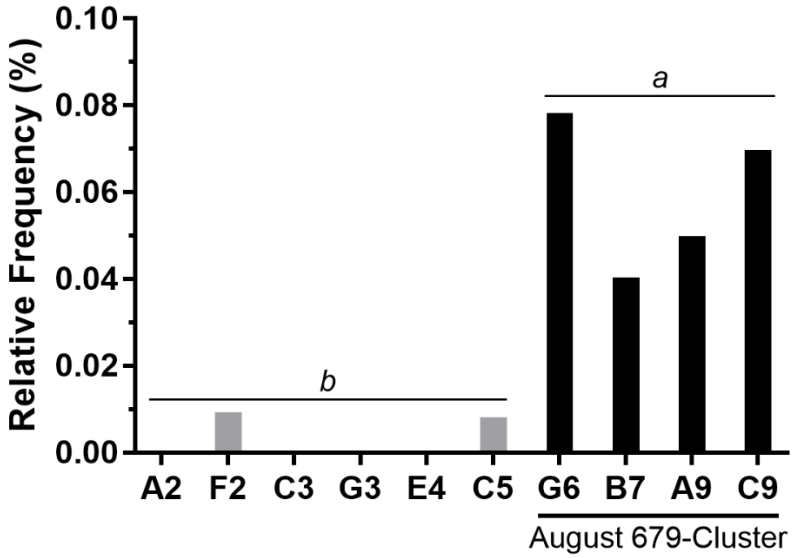
[c] Bolded values indicate significant difference between August 679-Cluster and non-August 679-Cluster ($P<0.05$) as assessed through an analysis of variance (ANOVA) test with post-hoc pairwise Holm-Sidak multiple comparisons

[d] Full classification indicates genus-level classification to the 'Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium' group

Compared to other metagenomes assessed in August 2018, the August 679-Cluster had significantly greater relative proportions (between 1.94% and 4.38%) of sequences classified to *Rhodobacteraceae*, *Fibrobacter*, *Bacteroides ihuae*, *Rhizobiaceaee*, *Microbacteriaceae*, *Arcobacter butzleri*, *Pseudomonas*, *Pseudomonadaceae*, and *Elusimicrobium* (Table 4.2). Of these, all except *Fibrobacter,*

Bacteroides ihuae, and *Elusimicrobium* had sequence variants predicted by PICRUSt2 as contributing to toluene degradation capacity in metagenomic samples (Figure 4.4). As expected when using PICRUSt2, not all (though most) sequence variants within these phylotypes were predicted to have toluene-degrading functional capacity; however, this does not preclude potential functional capacity for those organisms (Garza and Dutilh 2015; Langille 2018). Enrichment of specific taxa in contaminant-exposed microbiomes offers a line of evidence for their involvement in remediation activities (Winderl et al. 2008; Crombie et al. 2018; Roy et al. 2018). Therefore, toluene utilization by the phylotypes presented in Table 4.2 merits further investigation. Identification, isolation, and characterization of poplar endophytes capable of aromatic hydrocarbon degradation is of interest for future phytoremediation applications, as they might have potential as inoculants to enhance phytoremediation activity (Moore et al. 2006; Jambon et al. 2018).

A)



B)

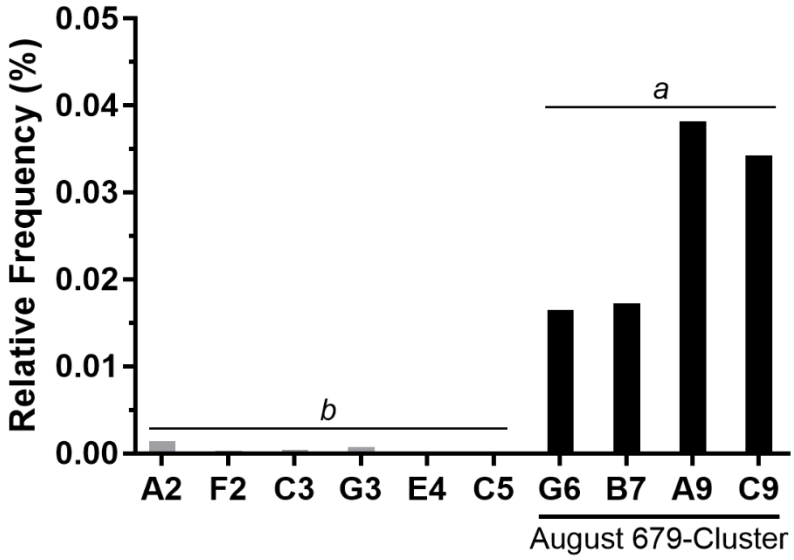


Figure 4.4: Relative frequency (%) of aerobic toluene degradation functional features in August 2018 stem tissue metagenomic samples via the *p*-cresol (A) and catechol (B) intermediate pathways predicted using PICRUSt2 metagenomic functional modeling. Dissimilar lowercase letters denote significant (P<0.05) differences in average relative frequency of functional features between 679- Cluster and non-679 Cluster samples.

4.5 Conclusions

This study explored the influence of *in planta* aromatic hydrocarbon exposure on hybrid poplar stem phyllosphere ecological structure at a field-scale application. Stem phyllosphere bacterial community structure was assessed in early spring and late summer under active phytoextraction conditions. The results indicated a seasonal change occurred in poplar stem microbiomes from the early to late growing season. A site-wide increase in stem microbiome diversity and richness was observed over time, coinciding with a marked decrease in abundance of a population of suspected pathogenic bacteria from the *Enterobacteriaceae* family. Generally, Gammaproteobacteria, Actinobacteria, Bacteroidia, Clostridia, Bacilli, Coriobacteriia, and Alphaproteobacteria were found in high relative proportions in stem phyllosphere communities at this site.

A substantial finding of this work was that distinct bacterial communities were selected for in trees growing in the more heavily impacted area of the site in late summer. These trees had the greatest measured concentrations of toluene in their transpiration streams, reflecting substantially greater groundwater toluene concentrations there compared to upgradient locations. Exposure to toluene in these trees appeared to cause a general increase in diversity and richness, beyond site-wide seasonal effects. These microbiomes were enriched in late summer with *Rhodobacteraceae*, *Fibrobacter*, *Bacteroides ihuae*, *Rhizobiaceaee*, *Microbacteriaceae*, *Arcobacter butzleri*, *Pseudomonas*, *Pseudomonadaceae*, and *Elusimicrobium* compared to stem microbiomes of other trees. Moreover, toluene-exposed stem microbiomes also had a greater predicted functional capacity for aerobic toluene degradation.

Together, these results improve the general understanding of hybrid poplar phyllosphere ecology in aromatic hydrocarbon phytoremediation systems, and autochthonous degrader communities therein. The present work demonstrates, on the field scale, how stem phyllosphere ecology can provide a line of evidence for contaminant biodegradation activity, as is common for other ecological niches. Moreover, it demonstrates further how the combination of taxonomic profiling with spatially-resolved subsurface contaminant concentration profiles can help refine site conceptual models for more comprehensive understanding of mechanistic phytoremedial processes.

Acknowledgements:

This work was supported by BP, the University Consortium for Field-Focused Groundwater Contamination Research, and the Natural Sciences and Engineering Research Council of Canada (NSERC). Special thank-you to Alan Scheibner (BP Canada) for ongoing support and contributions to the project. Thank-you to Kamini Khosla, Jonathan Gaiero, Micaela Tosi, Jemaneh Habtewold, Dasiel Obregón Alvarez, Steve Chapman, Maria Gorecka, Patryk Quinn, James Hommerson, Brent Coleman, Kevin B. Ecott, Sean Jordan, Ian Renaud, and Steve Wilson for technical support.

5 Degrader bacteria and fungi enriched in rhizosphere soil from a toluene phytoremediation site identified using DNA stable isotope probing

This manuscript is in preparation

Keywords: high-throughput sequencing; PICRUSt; metagenomic molecular biological tools; DNA-SIP; hybrid poplar; BTEX aromatic hydrocarbons

5.1 Abstract

Improved knowledge of the identity and ecology of contaminant-degrading organisms is paramount for effective assessment, monitoring, and remediation of aromatic hydrocarbon-impacted sites. In the present study, DNA stable isotope probing was used to identify autochthonous degraders in rhizosphere soil from an active hybrid poplar-based phytoremediation system incubated under semi field-simulated conditions. High-throughput sequencing of bacterial 16S *rRNA* and fungal ITS *rRNA* genes was performed on density-resolved metagenomic samples. Putative toluene-degrading bacteria were identified mainly within the Actinobacteria and Proteobacteria phyla. Predominant sequences within these groups classified as *Cupriavidus*, *Rhodococcus*, *Luteimonas*, *Burkholderiaceae*, *Azoarcus*, *Cellulomonadaceae*, and *Pseudomonas* organisms. *Purpureocillium lilacinum*, a well-known fungal toluene degrader, and *Mortierella alpina*, a fungal species not previously known to degrade toluene, were found to actively assimilate toluene as well. Several strains of *Mortierella elongatus*, a fungal poplar endophyte, were also implicated as potential toluene degraders in this study. This study made use of PICRUSt2 predictive taxonomic functional modeling to characterize density-resolved DNA samples and support successful isolation of stable isotope-labeled DNA.

Finally, four unique bacterial sequences from putative degraders, classified within the *Bdellovibrionaceae* (OM27 clade), *Intrasporangiaceae*, or *Chitinophagaceae* families, or within the Sphingobacteriales (AKYH767) order were not found within PICRUSt2-generated metagenomes and represent potentially novel putative toluene-degrading species. This study illustrates the power of combining stable isotope amendment with advanced metagenomic and bioinformatic techniques to link biodegradation activity with unisolated microorganisms.

5.2 Introduction

Aromatic hydrocarbons are common environmental contaminants in soil and groundwater, often deriving from anthropogenic releases of petroleum and petroleum derivatives. Though toxic and mobile, aromatic hydrocarbons, including benzene, toluene, ethylbenzene, and xylene isomers (BTEX) are not substantially recalcitrant and are strongly amenable to microbe-mediated attenuation (Davee and Sanders 2000; WHO 2004; Thierrin et al. 2009). In management of BTEX-impacted sites, detection and quantification of known degraders serves as a valuable indicator for biodegradation activities and is often a prerequisite for remediation activities (Kao et al. 2010; Larentis et al. 2013; Spini et al. 2018). To this end, a firm understanding of the microbial ecology of degraders at contaminant impacted sites has enabled site managers to demonstrate occurrence of active biodegradation and to help meet regulatory and site management goals (Cho et al. 2013; Key et al. 2014; BenIsrael et al. 2019).

Increasingly affordable, high-throughput metagenomic molecular biotechnological tools are making ecological information of this nature more accessible and powerful for site assessments (Techtmann and Hazen 2016). This includes bioinformatic predictive functional profiling tools, such as PICRUSt2 (Douglas et al. 2019), which effectively deduces functional capabilities of diverse, uncultured communities using only taxonomic marker genes. This tool (including earlier iterations) has increasingly been employed to

characterize contaminant degrader communities and inform on their ecological dynamics (Mukherjee et al. 2017; Roy et al. 2018; Julio et al. 2019). However, taxonomic functional modeling relies on reference databases, often overrepresenting certain taxa or missing undiscovered organisms (Garza and Dutilh 2015), and excludes sequences from analysis that are not sufficiently similar to reference genome sequences. Therefore, increasing the breadth of reference genomes of known aromatic hydrocarbon degraders is essential to increase the relative efficacy of predictive taxonomic functional modeling.

Also underrepresented in the literature are fungal organisms that can degrade aromatic hydrocarbons. Known fungi can transform and biodegrade (typically as co-metabolic processes) these compounds, while rare fungal isolates, particularly in the *Cladophialophora* and *Exophiala* genera, have been found that can use them (specifically toluene) as a carbon and energy source (Parales et al. 2008). The most benefit, from a remediation standpoint, comes from fungi such as these that can perform full mineralization of contaminants (Prenafeta-Boldu et al. 2006). In the field of phytoremediation (i.e. plant-derived attenuation of environmental contaminants in impacted media), isolation of these and bacterial degraders than can colonize plant tissues is of particular interest for enhancing remediation systems (Jambon et al. 2018).

Overall, there is a motivation to broaden the collection of known degraders to improve diagnostic tools and to better understand the function and community dynamics at impacted sites. Studying substrate utilization by microbial communities, however, is challenging. Limitations in culture-dependent techniques to study microbial communities are well established (Stefani et al. 2015), and enrichment of organisms in a mixed culture cannot alone be used as conclusive evidence for active contaminant utilization (Wolicka et al. 2009). DNA stable isotope probing (DNA-SIP) is a powerful emerging molecular technique that can reliably overcome these limitations for studying degraders. In DNA-SIP, utilization of a stable isotope-labeled substrate by degraders can be used to definitively implicate them in active biodegradation (Neufeld et al. 2007). The method

makes use of the increased buoyant density of nucleic acids synthesized using assimilated ^{13}C from metabolized labeled compounds, whereby ^{13}C-labeled DNA can be isolated and characterized to target investigation of degraders. To isolate ^{13}C-labeled DNA, extracted DNA is ultracentrifuged in a CsCl solution, which forms a gradient of buoyant densities; heavy, ^{13}C-labeled DNA will migrate to a different buoyant density (isopycnic point) than lighter, unlabeled DNA. The centrifuged DNA is then separated into multiple density fractions to isolate ^{13}C-labeled DNA for further analysis.

The present study sought to identify known and candidate toluene-degrading bacteria and fungi from an impacted field site to expand the knowledge of active toluene-degrading species. Rhizosphere soil from an active toluene phytoremediation system was used in a DNA-SIP incubation to generate stable isotope-labeled metagenomic samples. High-throughput DNA sequencing was applied to characterize density-resolved DNA samples and to taxonomically characterize inferred toluene degraders. The present study also made use of predictive taxonomic functional modeling to characterize density-resolved DNA samples and to support identification of toluene-degrading organisms. Results of this study offer a reference library of bacterial and fugal taxonomic marker genes implicated in toluene mineralization.

5.3 Materials and Methods

5.3.1 Sample collection and microcosm toluene incubations

Toluene incubations were carried out in nine microcosms constructed using 60 mL serum bottles with rubber stoppers and aluminum crimp seals (Table 5.1). Rhizosphere soil previously in association with hybrid poplar (*Populus deltoides* × *nigra* OP-367) roots at a toluene-impacted phytoremediation field site was used as an inoculum source. The soil was obtained in November 2016 over a depth range of 0.97 to 1.93 m below ground surface, in a bioactive toluene degradation zone (BenIsrael et al. 2019). Briefly, poplar root material was obtained at multiple locations using a CAT® 420 excavator with an 18-

inch bucket. Rhizosphere soil was separated from roots immediately and preserved in the dark at 4 °C for 16 months until microcosm construction. Soil was homogenized and sieved using a 2 mm mesh number 10 sieve (Fischer Scientific, Ottawa, Canada), then 3 g of soil were added to each microcosm pre-charged with 25 mL of growth media with trace metals, composed of NaCl (0.01 g/L), $CaCl_2$ (0.05 g/L), NH_4Cl (0.2 g/L), $K_2HPO_4 \cdot 3H_2O$ (0.1 g/L), KH_2PO_4 (0.055 g/L), $MgCl_2 \cdot 6H_2O$ (0.2 g/L), $Na_2S \cdot 9H_2O$ (0.5 g/L), $FeCl_2 \cdot 4H_2O$ (0.1 g/L), $MnCl_2 \cdot 4H_2O$ (0.001 g/L), $CoCl_2 \cdot 6H_2O$ (0.0017 g/L), $ZnCl_2$ (0.002 g/L), H_3BO_4 (0.00019 g/L), $NiCl_2 \cdot 6H_2O$ (0.0005 g/L), $Na_2MoO_4 \cdot 2H_2O$ (0.0002 g/L), $CuCl_2 \cdot 2H_2O$ (0.00039 g/L), and KNO_3 (1 g/L), $MgSO_4$ (1 g/L). Prior to incubation with toluene, microcosms were sealed and stored at room temperature for a starvation period of 15 weeks to reduce available carbon sources. At the start of the incubation (day 0), cAMP (1.646 g/L), 0.15 g Amberlite™ XAD7 (as an adsorbent resin), and 3 µL of ^{12}C or ^{13}C-labeled (Cat. No. 606618, Sigma-Aldrich Co., Missouri, USA) toluene were added to carbon-starved microcosms, according to Table 5.1. Control microcosms were constructed using sterile autoclaved soil for use in headspace gas analyses. An unamended control was constructed to assess pre-toluene-incubation community profiles. Microcosm bottles were incubated in the dark at 20 °C with shaking at 175 rpm for up to 18 days. The incubation temperature and inclusion of mineral growth media in these incubations was selected to partially simulate field-relevant temperature and toluene exposure conditions. These were based on historical shallow groundwater temperature measured during mid-growing season at the phytoremediation site from which soil inoculum was obtained (Fernandes 2017) and to reflect the exposure to toluene-impacted groundwater that these soils experience under seasonal high water table conditions on site (BenIsrael et al. 2019).

Table 5.1: Microcosm experimental set-up for toluene amendment, headspace CO_2/toluene GC-MS analysis and DNA extraction days, and DNA concentrations yielded.

Bottle[a]	Sample name	Toluene (3 µL)	GC-MS	DNA extraction	DNA concentration (ng/µL)
1	^{12}C Day 1	^{12}C	Day 1	-	
2	^{12}C Day 13	^{12}C	Day 13	Day 13	14.99
3	^{12}C Day 18	^{12}C	Day 18	Day 18	43.26
4	^{13}C Day 1	^{13}C	Day 1	-	
5	^{13}C Day 13	^{13}C	Day 13	Day 13	24.06
6	^{13}C Day 18	^{13}C	Day 18	Day 18	55.24
7 – sterile	Sterile Day 1	^{12}C	Day 1	-	
8 – sterile	Sterile Day 13	^{12}C	Day 13	-	
9 – sterile	Sterile Day 18	^{12}C	Day 18	-	
10 – unamended	Unamended	-	-	Day 1	4.63

[a]Autoclaved soil used in microcosms 7–9 as a sterile control.

5.3.2 Toluene and CO2 headspace sampling

Toluene and CO_2 relative content in microcosm headspaces were assessed on days 1, 13, and 18 (Table 5.1). Toluene was assessed by analyzing duplicate headspace gas samples using a gas chromatograph interfaced with a mass spectrometer (GC-MS) at the AAC Mass Spec Facility (University of Guelph, Guelph, Canada). The GC-MS system (Agilent, Santa Clara, CA, USA) included an HP 7890A GC interfaced with an Agilent 5975C mass selective single quadrupole detector. Chromatography was performed using an HP-5MS (30 m x 0.25 mm x 0.5 µm 5%-phenyl-methylpolysiloxane) stationary phase column (Agilent Technologies, Santa Clara, California, USA). The oven temperature was held at 40 °C for 2 min and raised from 40 to 60 °C at 10 °C/min, then from 60 to 230 °C at 120 °C/min, for a total runtime of 5.4 min. Helium (grade 5.0) was used as the carrier gas and the flow through the column was held constant at 1 mL/min. The injection volume was 100 µL with a gas tight syringe in the splitless mode. The injection temperature was 220 °C and the transfer line temperature was 260 °C. The GC column was equilibrated for 5 min prior to each analysis. The MS was operated in scan mode (mass range 10-150 amu) at 5 scans/sec. The MS source temperature was set at 230 °C and the MS quadrupole temperature was 150 °C. The data was acquired in

electron impact (EI) positive ionization mode using 70eV energy and the data analyzed on Agilent MSD Chemstation software (version E02.02.1431). CO_2 content was assessed by analyzing triplicate 5 mL headspace gas samples using an SRI 8610C GC (SRI Instruments, California, USA) with a 6 ft Haysep D packed column and a flame ionization detector (FID) with a methanizer, nitrogen as a carrier gas, and a 20 µL sample loop. Relative toluene and CO_2 content in headspace samples were calculated by integrating chromatogram peak areas.

5.3.3 DNA extraction, isopycnic centrifugation, and density-discrete fractionation

Microcosm sediment was collected for DNA extraction by centrifugation at 4,500 rpm for 3 min at room temperature (Table 5.1). DNA was extracted from triplicate 0.5 g wet sediment samples (1.06 g total on average on a dry weight basis) using a Powersoil DNA Isolation Kit (cat no. 12888-100, Mo Bio Laboratories, Inc., Carlsbad, CA USA) and pooled. DNA samples were stored at -20 °C until further processing and additional sediment samples were used to evaluate gravimetric moisture content. Isopycnic centrifugation, fractionation, and DNA precipitation were carried out according to methods described in Dunford and Neufeld (2010). Briefly, centrifugation was carried out using an Optima™ L-100 XP Ultracentrifuge with a VTi 65.2 Vertical-Tube Rotor (Beckman Coulter Inc., Brea, CA, USA) at 44,100 rpm for 40 hours, at a temperature of 20 °C. Fractionation was carried out using a BSP Single Speed Syringe Pump (Braintree Scientific Inc., Braintree, MA, USA). Buoyant densities (BDs) of fractions were evaluated using an AR200 Automatic Digital Refractometer (Ametek Reichert Technologies, Depew, NY, USA) using an empirically derived refractive index curve. Afterwards, DNA precipitation was carried out using linear polyacrylamide as a carrier for precipitation in a polyethylene glycol solution.

5.3.4 DGGE assessment of density-resolved fractions

Denaturing gradient gel electrophoresis (DGGE) was performed on 16S *rRNA* genes amplified from density-resolved fractions according to methods described in Piterina and Pembroke (2013). Briefly, PCR was carried out using the GC-338F (5'-CGCCCGCCGCGCGCGGCGGGCGGGGCGGGGGCACGGGGGGACTCCTACGGGAGGCAGCAG) and 518R (5'-ATTACCGCGGCTGCTGG) primers in 25 µL reactions containing 12.5 µL GoTaq® Green Master Mix (Promega Corp., Madison, WI, USA), 1.25 µL of each primer (10 pmol/µL), and 1 µL of template. Reactions were carried out in a T100™ Thermal Cycler (Bio-Rad Laboratories Inc., Hercules, CA, USA) with an initial denaturation for 3 min at 95 °C, 36 cycles of denaturation (30 s at 95 °C), annealing (30 s at 55 °C), and extension (30 s at 72 °C), and a final extension for 5 min at 72 °C. DGGE was carried out at 60 V over 16 hours at 60 °C using a DCode System (Bio-Rad Laboratories Inc., Hercules, CA, USA) in 10% gels containing a 30% to 80% denaturing gradient. Gels were visualized after incubation with ethidium bromide to compare community profiles between BD fractions.

5.3.5 DNA sequencing, bioinformatics, and statistical analyses

DNA sequencing was performed targeting portions of the V4 region of bacterial/archaeal 16S *rRNA* genes and portions of the ITS1 region of fungal *rRNA* genes in whole-microcosm DNA samples and in select BD fractions within four BD ranges: 1.704–1.710, 1.72–1.734, 1.751–1.753, and 1.758–1.762 g/mL. Heavy ^{13}C-labeled DNA is typically found in fractions within a BD range of 1.720 to 1.735 g/L when using the SIP protocol employed here (Dunford and Neufeld 2010); these fractions were designated in this study as the 'labeled BD fractions'. Heavier fractions (1.751–1.753 and 1.758–1.762 g/mL) were therefore designated as 'heavy unlabeled fractions' and lighter fractions (1.704–1.710 g/mL) as 'light unlabeled fractions'. Amplicon preparation and sequencing were conducted at the McGill University and Génome Québec Innovation Centre (Montréal, QC, Canada) using the Illumina MiSeq PE300 amplicon sequencing platform.

The 515F (5’-GTGYCAGCMGCCGCGGTAA) and 806R (5’-GGACTACNVGGGTWTCTAAT) primer set was used to amplify 16S *rRNA* genes, while the ITS1F (5’- CTTGGTCATTTAGAGGAAGTAA) and ITS2 (5’-GCTGCGTTCTTCATCGATGC) primer set was used for fungal *rRNA* genes, as described in the Earth Microbiome Project protocol (Caporaso et al. 2012).

Raw FASTQ sequence files for 16S and ITS *rRNA* genes were processed and analyzed using Qiime2 v. 2019.1 software (Bolyen et al. 2019). For 16S sequences, forward and reverse reads of each sample were merged using the demultiplexing protocol for single-end reads barcoded according to the Earth Microbiome Project protocol (Caporaso et al. 2012). Sequence quality control was performed using the DADA2 pipeline for detecting and correcting sequencing errors (Callahan et al. 2016): a trim length of 25 (forward and reverse) and truncation lengths of 220 (forward) and 206 (reverse) were selected based on sequencing quality scores. Sequences were grouped into sequence variants based on 100% similarity. Sequence taxonomy was then classified using the Silva 132 classifier for 99% classification targeting the 515–806 region with seven-level taxonomy (Quast et al. 2013). The classifier used reflected the recent reclassification of Betaproteobacteria, which contains common toluene-degrading organisms (Parales 2010), as a new order, Betaproteobacteriales, within the Gammaproteobacteria class (Parks et al. 2018). Sequences were filtered to remove non-bacterial/archaeal, mitochondrial, and chloroplast sequences, rarefied to the smallest number of observed reads (61,178 reads for whole-microcosm samples and 36,205 for BD fractions), then the mafft, mask, and FastTree protocols were used to generate rooted and unrooted phylogenetic trees of aligned representative sequences for use in diversity analyses. Sample diversity metrics were generated for α-diversity, including Shannon Diversity Index, Pielou Evenness, Faith Phylogenetic Diversity, and number of observed sequence variants. The National Center for Biotechnology Information (NCBI) nucleotide Basic Local Alignment Search Tool (BLASTn) tool was used to evaluate evolutionary relationships of detected sequences to reference sequences/organisms.

ITS sequences were similarly processed, with the following modifications. Merged sequences were trimmed using the ITSxpress pipeline (Rivers et al. 2018) before controlling for quality using DADA2, in which no trimming or truncation was performed. Grouped sequence variants were taxonomically classified using the UNITE v.8.0 Dynamic Classifier for fungi (UNITE Community 2019). No taxonomic filtering was performed, and rarefaction prior to diversity analysis was performed to 11,333 reads for whole microcosm samples and 17,306 for BD fractions.

To identify putative degraders, organisms whose sequences demonstrated at least a threefold increase (Wilhelm et al. 2018) in average relative abundance of 16S or ITS *rRNA* genes in the ^{13}C-toluene labeled BD fractions relative to ^{12}C-toluene fractions of the same density range were deemed with high likelihood to have mineralized toluene. The comparison to ^{12}C-toluene samples serves as a control for artifacts of ultracentrifugation and variable GC content of nucleic acids, which can cause unlabeled nucleic acids to separate into the labeled BD fractions (Buckley et al. 2007). Where this response ratio was less than 3, or not able to be calculated (i.e. sequence not found in corresponding ^{12}C-toluene labeled BD fractions), substantial enrichment in corresponding toluene-amended microcosms provided evidence of toluene use.

Predictive evolutionary modeling of inferred community function in whole-microcosm samples and select BD fractions was performed using PICRUSt2 analysis (Douglas et al. 2019) of 16S *rRNA* genes. A Nearest Sequenced Taxon Index (NSTI) cut-off value of 2 was used. The MetaCyc superpathway of aerobic toluene degradation (PWY-5183) functional feature (Caspi et al. 2014) was targeted to assess predicted capacity for aerobic toluene degradation, and functional capacity is reported as the frequency of this functional feature relative to all predicted functional features. Differences in relative functional feature frequency between samples were evaluated using STAMP v. 2.1.3 software (Parks et al. 2014) using multiple group analysis of variance (ANOVA)

tests with Tukey-Kramer pairwise comparison. Statistical significance in all cases was determined using an α=0.05 threshold.

Putative degrader bacterial 16S and fungal ITS *rRNA* sequences have been submitted to GenBank with the accession numbers MN909835–MN909862 and MN910273–MN910277, respectively.

5.4 Results and discussion

5.4.1 Toluene utilization during incubation

Rhizosphere soil from the poplar-based toluene phytoremediation system was incubated in microcosms with growth media and ^{13}C-labeled and unlabeled toluene to identify toluene-degrading microorganisms under semi-field-simulated conditions. Toluene and CO_2 content in microcosm headspace gas were evaluated on days 1, 13, and 18 of incubation to assess microbial activity and toluene utilization. Both functions were evidenced by decreasing toluene and increasing CO_2 content in headspace gas of non-sterile toluene-amended microcosms over time, while quantities of both gases remained relatively constant in sterile-soil microcosms (Figure 5.1). In non-sterile microcosms, relative toluene content dropped by approximately 3 orders of magnitude by day 13 in both ^{13}C and ^{12}C incubations and was not quantifiable in the ^{13}C microcosm on day 18. CO_2 content correspondingly increased by approximately one order of magnitude. Both CO_2 and toluene content remained relatively constant in sterile soil microcosms, indicating that loss of toluene was a result of toluene metabolism in non-sterile microcosms. Concentrations of DNA extracted from non-sterile toluene-amended microcosms on days 13 and 18 increased by an order of magnitude compared to DNA extracted from the unamended control (day 1) and approximately doubled from day 13 to day 18 (Table 5.1). This increase in biomass further supports that microbial communities were active in non-sterile microcosms over the incubation period.

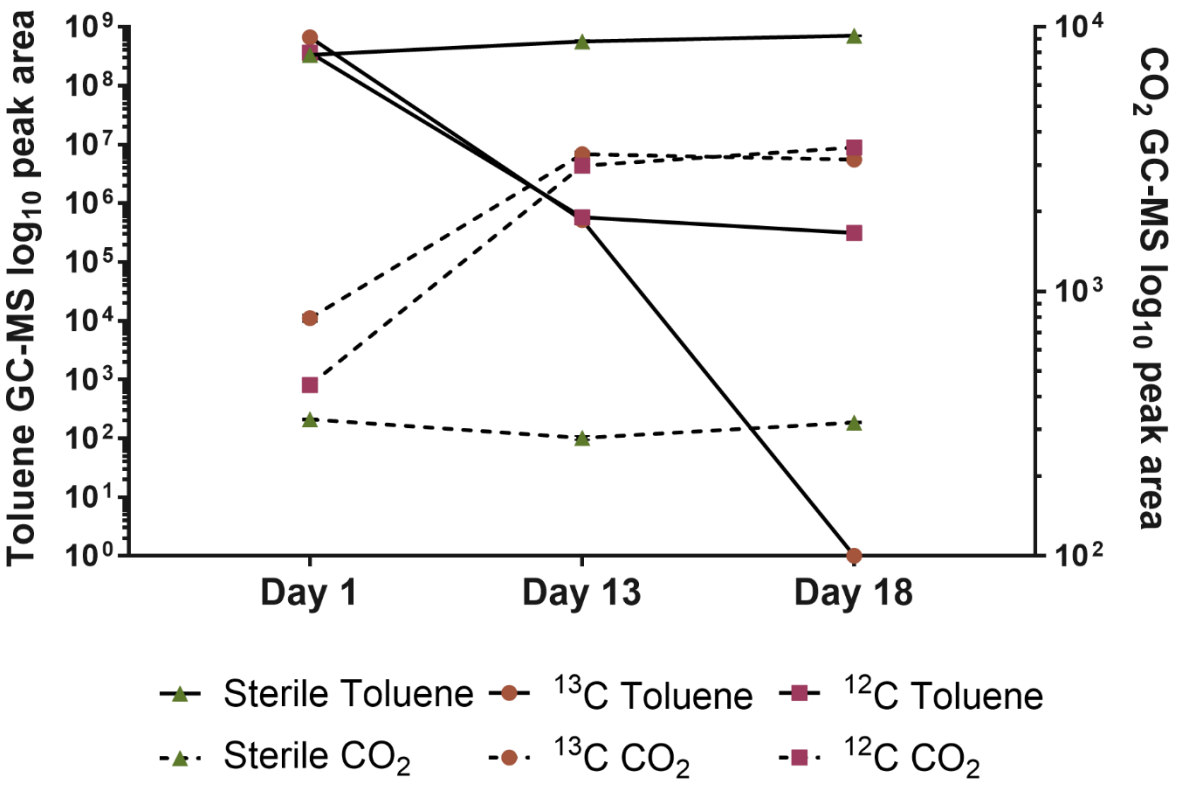


Figure 5.1: Toluene and CO_2 amounts in microcosm headspace. Values are log peak area of GC chromatograms (±SD). Toluene (n = 2) CO_2 (n = 3). Note: error bars are omitted where deviations are too small to visualize.

5.4.2 Community profiles in whole microcosms

Whole-community-level responses to toluene amendment were assessed by high-throughput DNA sequencing of bacterial/archaeal 16S and fungal ITS *rRNA* genes in whole-microcosm DNA samples. Most 16S *rRNA* sequence variants were classified as bacteria, while 0.068%–1.784% classified as archaea. Toluene amendment was observed to cause a selection and enrichment in a subset of the original bacterial/archaeal community. Archaea were reduced after incubation – the highest proportion of archaea observed went from 1.784% (Unamended) to 0.286% of sequences after toluene amendment. Moreover, the number of observed sequence variants and

diversity indices decreased in the amended samples compared to the day 1 unamended control (Table 5.2). Shifts in diversity were more variable for fungal communities than for bacteria/archaea (Table 5.2). The number of observed ITS sequence variants were similar in amended samples compared to the unamended control on day 13, but noticeably decreased on day 18. Therefore, a subset of fungal organisms appeared to be selected for after toluene amendment, but not as quickly as bacteria/archaea. This is consistent with relatively slow growth of fungi in liquid, aromatic hydrocarbon-amended culture conditions (Prenafeta-Boldu et al. 2006).

Table 5.2: Diversity indices for bacterial/archaeal 16S and fungal ITS *rRNA* gene sequences in whole-microcosm samples.[a]

	Day	Sample	Shannon Index	Pielou Evenness	Faith Phylogenetic Diversity	Observed Sequence Variants
Bacteria	1	Unamended	9.35	0.9	80.16	1330
	13	^{13}C	6.06	0.69	41.1	459
		^{12}C	5.08	0.58	43.31	425
	18	^{13}C	4.9	0.57	41.37	405
		^{12}C	5.34	0.61	40.67	429
Fungi	1	Unamended	2.5	0.48	7.64	37
	13	^{13}C	4.18	0.78	7.7	42
		^{12}C	3.65	0.73	4.93	32
	18	^{13}C	3.18	0.65	5.21	29
		^{12}C	2.82	0.6	4.54	26

[a] $n = 1$ for all

Toluene amendment also altered the taxonomic composition of bacterial/archaeal and fungal communities. Before toluene incubation, bacterial/archaeal communities were composed on the class level predominantly of Thermoleophilia (10.78%), Acidobacteria Subgroup 6 (9.08%), Actinobacteria (7.56%), Alphaproteobacteria (6.90%), Gammaproteobacteria (5.91%), Rokubacteria NC10 (5.14%), and Gemmatimonadetes (4.77%) (Figure 5.2). After toluene incubation, sequences with the greatest relative abundances (averaged between all amended samples) belonged mainly to

Gammaproteobacteria (57.78%), Actinobacteria (21.83%), Alphaproteobacteria (5.07%), and Bacilli (4.90%) (Figure 5.2).

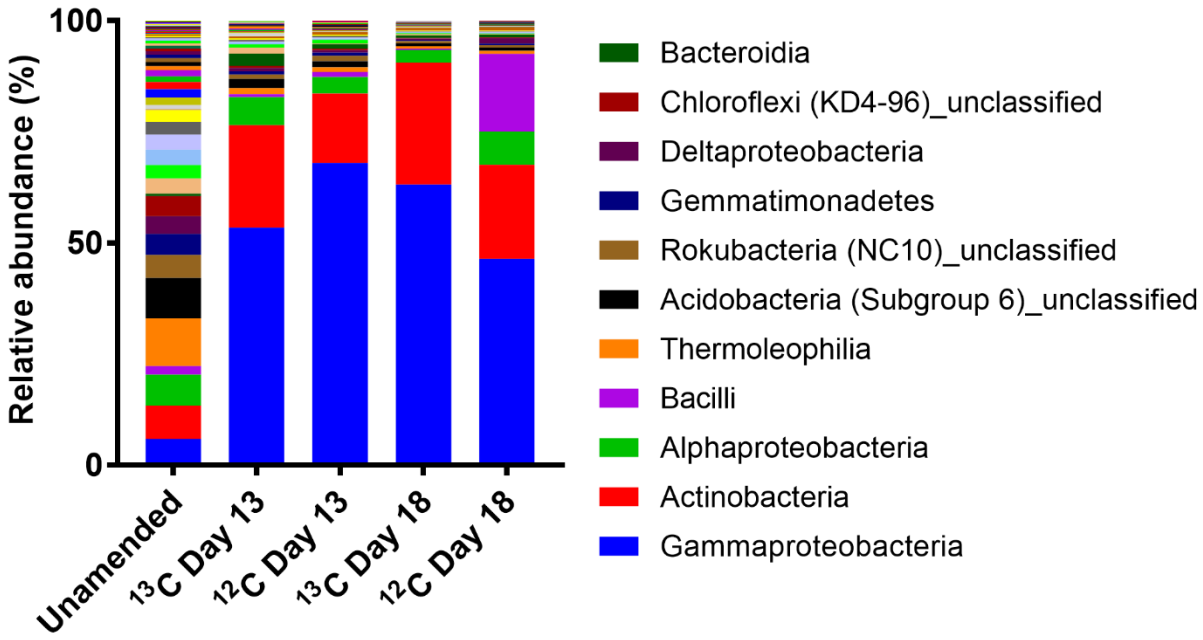


Figure 5.2: Relative bacterial/archaeal 16S *rRNA* sequence abundances as classified on the class level in whole-microcosm samples.

With fungal organisms, the unamended control was composed predominantly on the class level by Mortierellomycetes (93.19%) and Sordariomycetes (4.68%) (Figure 5.3A). After toluene incubation, the greatest average relative abundances between amended samples on the class level belonged to Sordariomycetes (51.21%), Mortierellomycetes (43.47%), and unclassified Rozellomycota (3.09%) (Figure 5.3A). On the species level, fungal communities went from being dominated by *Mortierella alpina* (92.86%) and *Metarhizium robertsii* (1.94%) in the unamended control, to being composed mostly (greatest relative abundances between amended samples) of *Mortierella alpina* (27.24%), unclassified Sordariomycetes (18.79%), *Mortierella elongata* (15.88%), *Fusarium solani* (13.12%), *Dactylonectria estremocensis (*7.74%), unclassified

Nectriaceae (3.20%), and unclassified Rozellomycota (3.09%) after toluene incubation (Figure 5.3B).

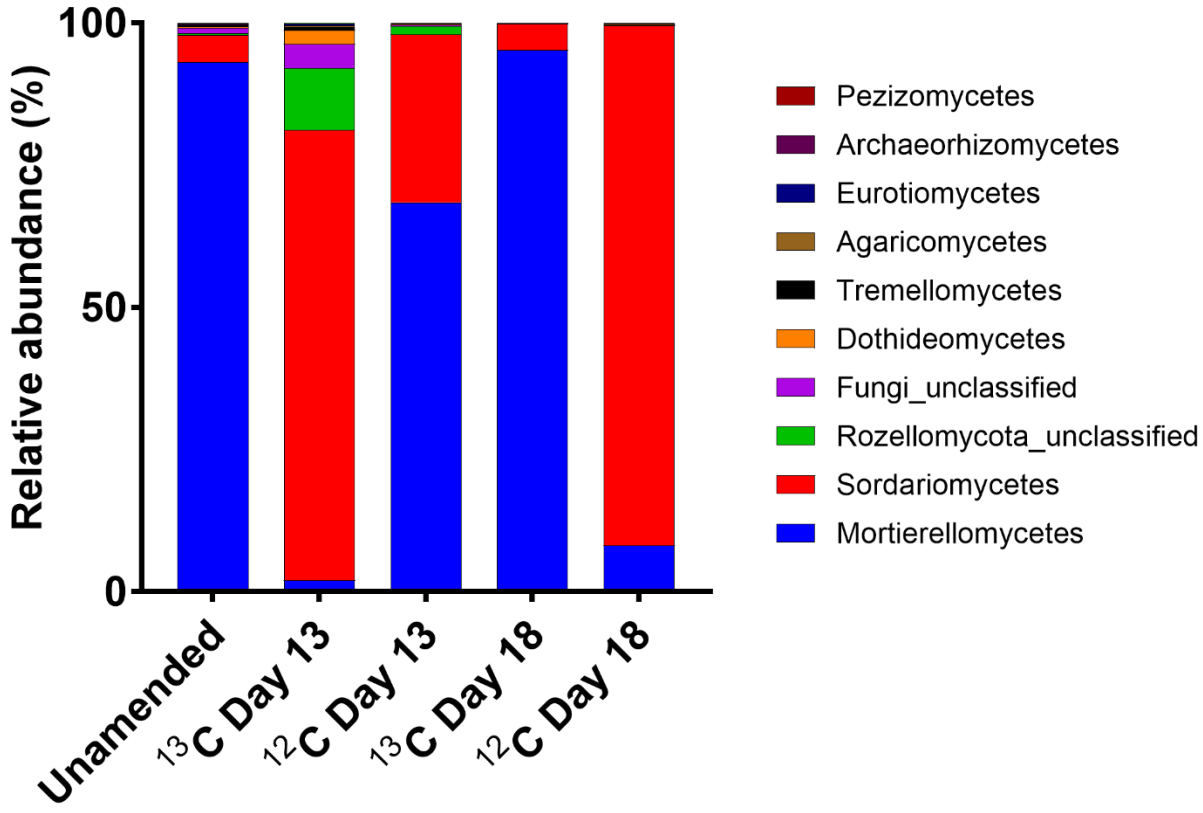


Figure 5.3: Relative fungal ITS *rRNA* sequence abundances as classified on the class level in whole-microcosm samples.

5.4.3 Community and functional profiles across buoyant density fractions

DNA from whole microcosm samples was separated into density-resolved fractions using isopycnic centrifugation to target assessment of ^{13}C-labeled DNA in labeled BD fractions, i.e., fractions within a buoyant density range expected to contain ^{13}C -labeled DNA (Dunford and Neufeld 2010). Preliminary screening of communities in BD fractions was performed using DGGE analysis of 16S *rRNA* genes. Community profiles observed provided evidence for unique communities in the labeled BD fractions

of the ^{13}C-toluene samples and therefore successful isolation of ^{13}C-labeled DNA from ^{13}C-toluene assimilators (Figure C.1). Community profiles in these fractions were notably distinct compared to other fractions for ^{13}C-toluene samples but were consistent between labeled and unlabeled BD fractions for ^{12}C-toluene samples. High-throughput sequencing of bacterial/archaeal 16S *rRNA* genes in select BD fractions further demonstrated this effect in ^{13}C-toluene samples (Figure 5.4). For day 13 this was most notable for bacteria in the genera *Rhodococcus* and *Azoarcus*, both of which are well known to contain degrader species (Zhou et al. 1997; Weathers et al. 2015). By day 18, a succession of *Pseudomonas*, increase in relative abundance of *Azoarcus*, and decreased abundance of *Rhodococcus* were observed, and labeled BD fractions were again distinct from other fractions (Figure 5.4B). For fungi, day 13 community profiles also demonstrated a shift in composition between the light, unlabeled BD and labeled BD fractions, particularly for sequences from *Mortierella alpina* (Figure 5.5A). Observation across BD fractions of different sequences classified as *Fusarium solani*, a species with known toluene-degrading strains (Morales et al. 2017), further supported successful isolation of ^{13}C-labeled DNA in the labeled BD fractions. *Fusariam solani* sequence variants in the labeled BD fractions represented 36.13% of sequences on average and were distinct from the other sequence variants classified as *F. solani* in the light, unlabeled BD fraction (Figure 5.5A). Sequences on day 18 in the light, unlabeled fraction and in one of the labeled BD fractions were not available for comparison, but labeled BD and other, heavier fractions showed dominance of *Mortierella alpina and Mortierella elongata* sequences (Figure 5.5B), which were dominant in the whole-microcosm community in general (Figure 5.3). Based on bacterial profiles resolved in these samples (Figure 5.4B), fungal DNA in the labeled BD fractions is believed to have been ^{13}C-labeled as well.

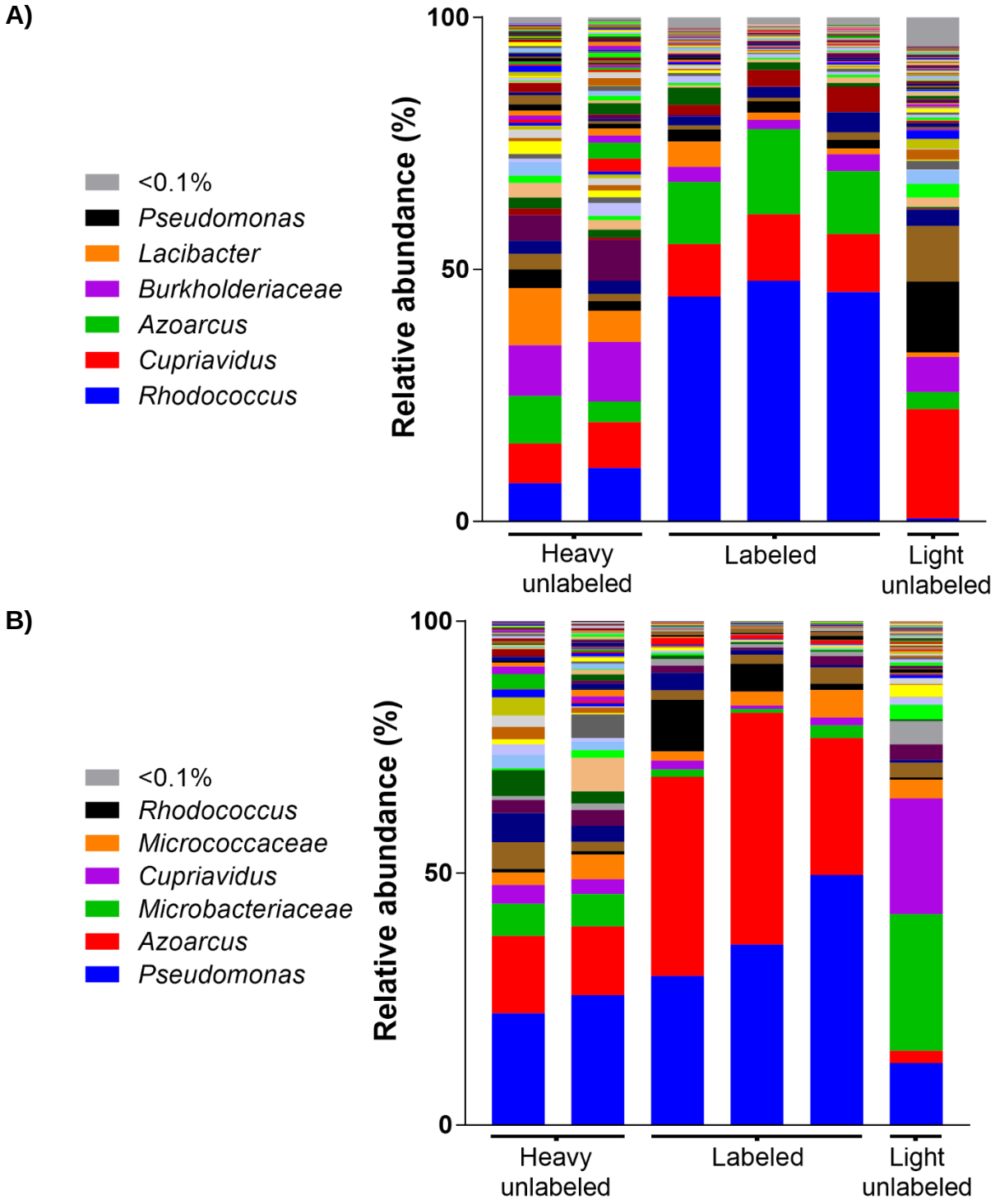
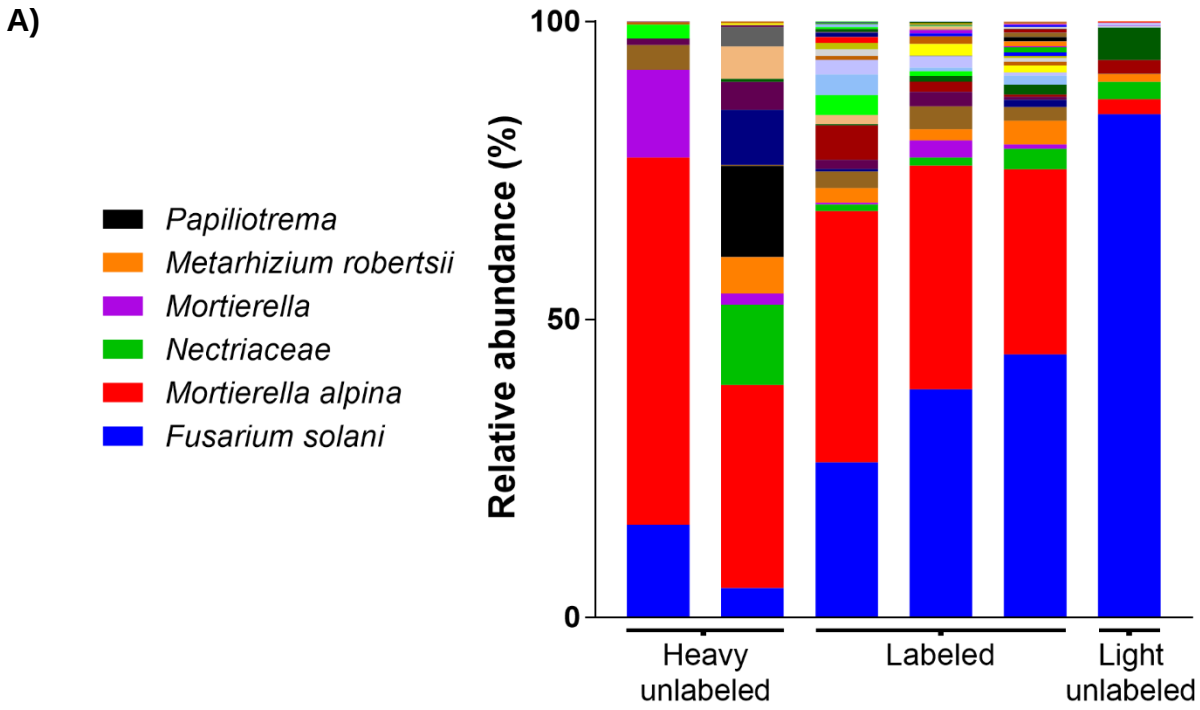


Figure 5.4: Bacterial/archaeal community profiles in BD fractions of ^{13}C-toluene samples for day 13 (A) and day 18 (B) generated from 16S *rRNA* gene sequences grouped based on common

taxonomy classification. Legends indicate highest classification (up to genus) of taxa in high relative abundance. BD fraction community profiles are presented from left to right with decreasing buoyant density over ranges of 1.762–1.758 and 1.753–1.751 g/mL (heavy unlabeled fractions), 1.735–1.720 g/mL (labeled fractions), and 1.710–1.704 g/mL (light unlabeled fractions).



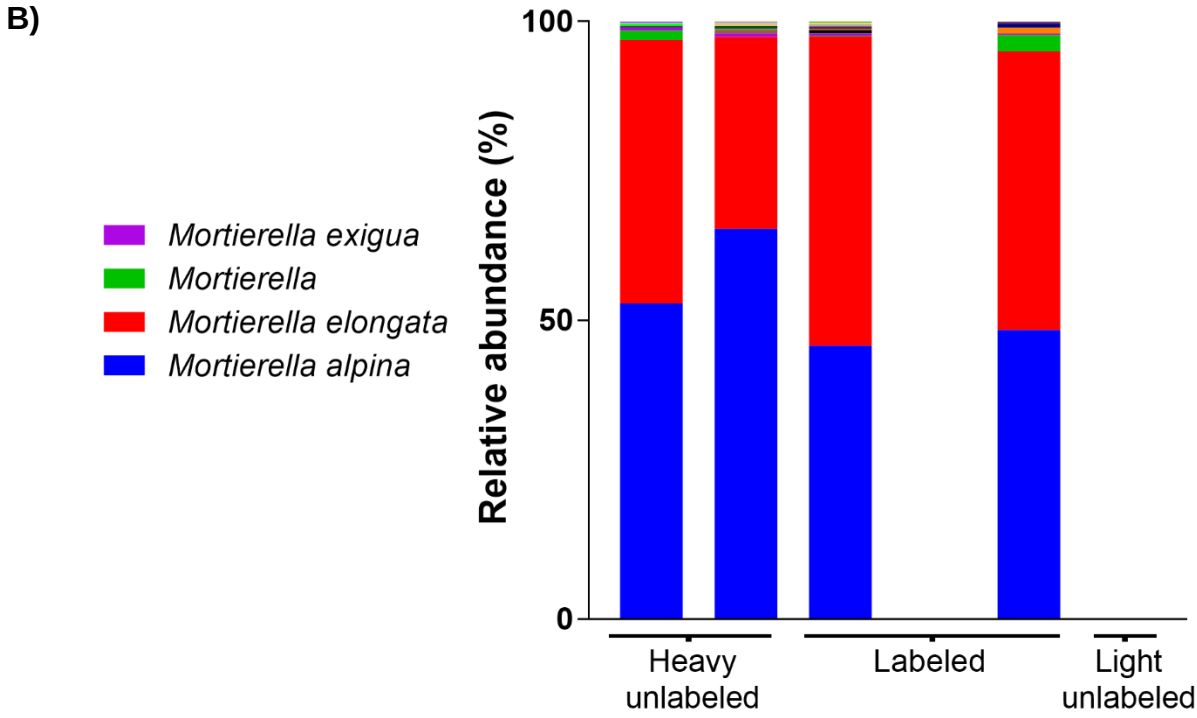


Figure 5.5: Fungal community profiles in BD fractions of ^{13}C-toluene samples for day 13 (A) and day 18 (B) generated from ITS *rRNA* gene sequences grouped based on common taxonomy classification. Legends indicate highest classification (up to species) of taxa in high relative abundance. BD fraction community profiles are presented from left to right with decreasing buoyant density over ranges of 1.762–1.758 and 1.753–1.751 g/mL (heavy unlabeled fractions), 1.735–1.720 g/mL (labeled fractions), and 1.710–1.704 g/mL (light unlabeled fractions). Note: sequences on day 18 in the light unlabeled and in one of the labeled BD fractions were not available for analysis.

Taxonomic functional modeling of sequenced 16S *rRNA* provided further evidence of ^{13}C-labeled DNA isolation in labeled BD fractions. Predicted capacity for toluene degradation was not significantly different ($P>0.05$) between ^{12}C-toluene and ^{13}C-toluene whole-microcosm samples; however, a significantly greater ($P<0.05$) capacity was predicted in the labeled BD fractions of ^{13}C samples than in ^{12}C labeled BD fractions (Figure 5.6A). Moreover, degradation capacity changed significantly ($P<0.05$) across the BD gradient in ^{13}C-toluene BD fraction samples, with the greatest predicted capacity

observed in the labeled BD range (Figure 5.6B). Together, the DGGE profiles, 16S/ITS community profiles, and the predicted functional profiles of ^{13}C-toluene and ^{12}C-toluene BD fractions support that isopycnic centrifugation successfully separated ^{13}C-labeled DNA into the expected BD fractions in the ^{13}C-toluene samples.

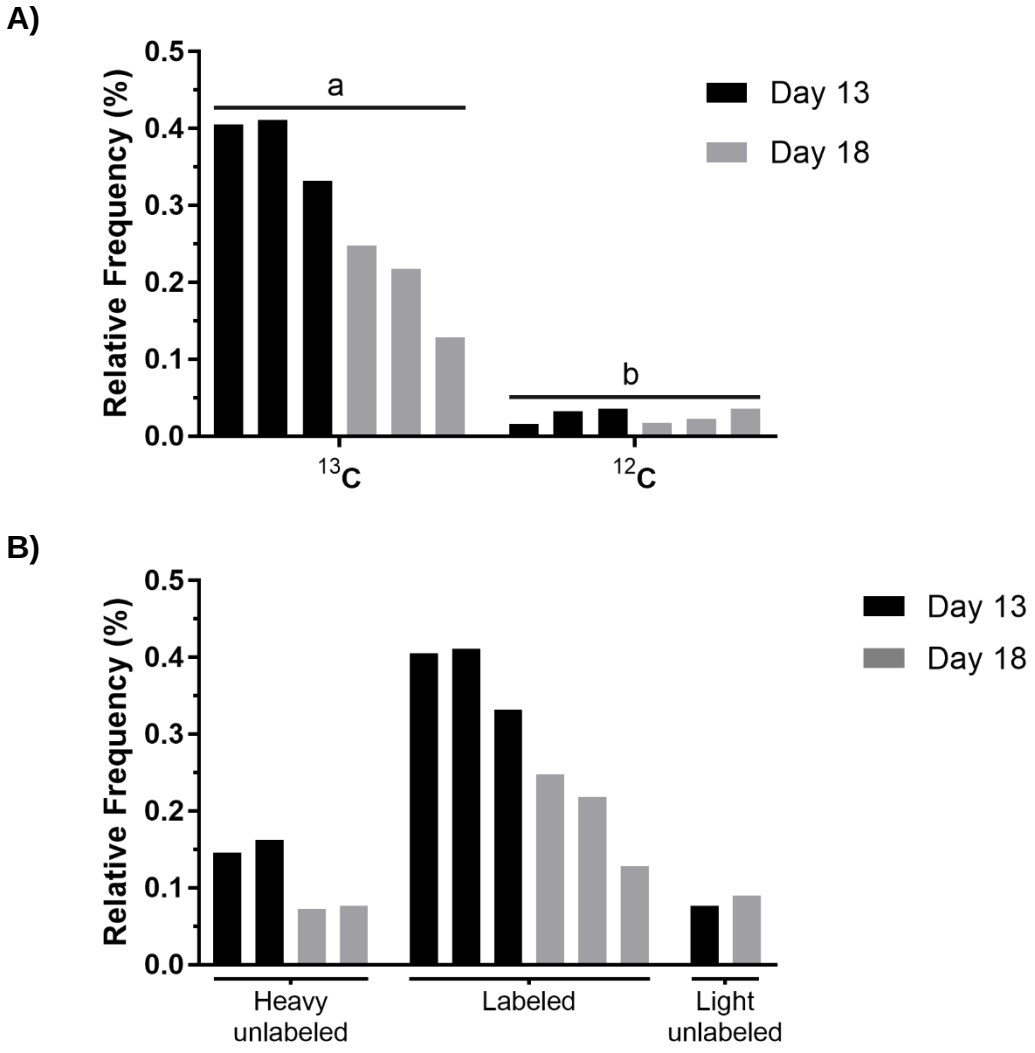


Figure 5.6: (A) Relative frequency (%) of the aerobic toluene degradation functional feature in labeled BD fractions (1.720–1.734 g/mL) predicted using PICRUSt2 metagenomic functional modeling for days 13 and 18 compared between ^{13}C- and ^{12}C-toluene samples. Lowercase letters denote significant ($P<0.05$) differences in average relative frequency of the aerobic toluene

degradation functional feature between labeled BD fractions from ^{13}C-toluene and ^{12}C-toluene-amended samples. (B) Relative frequency (%) of aerobic toluene degradation functional feature across a gradient of buoyant densities in ^{13}C-toluene samples predicted using PICRUSt2 metagenomic functional modeling. Note: values are plotted within each day with buoyant density decreasing from left to right, with ranges of 1.762–1.758 and 1.753–1.751 g/mL (heavy unlabeled fractions), 1.735–1.720 g/mL (labeled fractions), and 1.710–1.704 g/mL (light unlabeled fractions).

5.4.4 Putative bacterial toluene degraders

Putative toluene degraders were identified in this study based on enrichment in toluene-amended microcosms supported by inferred assimilation of ^{13}C. Assimilation of ^{13}C was identified by comparing bacterial and fungal profiles in labeled BD fractions between ^{13}C-toluene and ^{12}C-toluene samples. Organisms whose sequences demonstrated at least a threefold increase (Wilhelm et al. 2018) in relative abundance of 16S or ITS *rRNA* genes in the ^{13}C-toluene vs. ^{12}C-toluene labeled BD fractions were deemed with higher likelihood to have mineralized toluene (Table 5.3 and Table 5.4).

Inferred putative bacterial toluene degraders made up 44.43% and 31.25% of sequences in their respective whole-microcosm samples on days 13 and 18 and included sequences that were identical or very similar to know degraders (Table 5.3). Inferred putative degraders belonged almost exclusively to the Actinobacteria and Proteobacteria phyla, not surprisingly given the prevalence of aromatic degradation capacity in organisms within these two phyla in soil communities (Yang et al. 2014; Lima-Morales et al. 2016). No archaeal putative degraders were identified in this study based on ^{13}C/^{12}C-toluene response ratios, and indeed none were enriched compared to the Unamended control. Some archaea are known to biodegrade aromatic hydrocarbons in hypersaline and thermophilic environments (Krzmarzick et al. 2018). However, the biodegradative capacity of archaea in soil/aquifer environments, though not fully understood, is believed to be syntrophic in nature, such as the relationship between methanogenic archaea and

the fermentative or acetogenic bacteria that convert toluene to methanogenic precursors (Ficker et al. 1999). Therefore, failure to detect toluene-mineralizing archaea was not unexpected.

Dominant putative toluene-degrading organisms under incubation conditions of the present study were identified based on relative abundances of their sequences in the whole microcosm samples (Table 5.3). On day 13 these included sequences classified as *Cupriavidus* (two sequence variants, 15.62% and 4.7%), *Rhodococcus* (two sequence variants, 4.07% and 3.7%), *Luteimonas* (3.65%), *Burkholderiaceae* (2.49%), *Azoarcus* (2.4%), and *Cellulomonadaceae* (2.17%). Toluene-degrading organisms have previously been found in all of these taxa (Mechichi et al. 2002; Perez-Pantoja et al. 2008; Baquiran et al. 2012; Ortega-Gonzalez et al. 2013; Weathers et al. 2015; Lunsmann et al. 2016a). By day 18, dominant toluene-degrading bacteria still included *Cupriavidus* (12.64%) but exhibited a strong succession within the toluene-degrading consortium by *Pseudomonas* putative degraders (two sequence variants, 8.75% and 5%). Degraders in the Betaproteobacteriales order typically grow more slowly than Pseudomonadales degraders (Parales 2010), so it was somewhat surprising that *Cupriavidus*, a Betaproteobacteriales member, was able to outcompete *Pseudomonas*, a Pseudomonadales member, in the early stages of the incubation. This effect was likely due to competitive advantages of *Cupriavidus* under the physicochemical properties in the soil/culture media used, rather than superior degradative capacity (Palleroni et al. 2010). Indeed, sequences for both organisms were similar or identical to known toluene-degrading strains (Fahy et al. 2008; Perez-Pantoja et al. 2008).

Inferred putative bacterial toluene degraders identified in this study included apparently novel degrader phylotypes; however, none were substantially enriched by toluene amendment (i.e. at high relative abundance) under tested conditions. Bacterial 16S *rRNA* sequences of inferred degraders in labeled BD fractions of ^{13}C-toluene-amended samples were evaluated with PICRUSt2 to determine if this effect was

observable also on a predictive level. The model was able to predict aerobic toluene-degrading functional capacity in nearly all 16S *rRNA* sequences identified as belonging to putative degraders based on $^{13}C/^{12}C$-toluene response ratios, except for two sequences on day 13 and three sequences on day 18 (Table 5.3). These sequences classified within the *Bdellovibrionaceae* (OM27 clade), *Intrasporangiaceae*, or *Chitinophagaceae* families, or within the Sphingobacteriales (AKYH767) order. All came from rare organisms present at low abundance after incubation; they were below detection in the Unamended control and were only slightly enriched (below detection–0.22%) in their respective ^{13}C-toluene-amended microcosms (Table 5.3). The *Bdellovibrionaceae* family contains bacterial predators such as *Bdellovibrio* sp. (Williams et al. 2019); therefore, the possibility that *Bdellovibrionaceae* (OM27 clade) sequences might have been enriched in ^{13}C through predatory assimilation of ^{13}C-labeled biomolecules, rather than through direct toluene metabolism cannot be ignored. Though present at low abundance, the recovery of these sequences in the labeled BD fractions and their response ratios relative to ^{12}C-toluene controls suggest toluene mineralization activity by these organisms that merits further investigation. PICRUSt2 was, therefore, observed to be a robust tool for predicting the majority of sequences with putative toluene degradation capacity inferred through DNA-SIP under the incubation conditions of this study, and as a supporting line of evidence to evaluate more rare sequences from potentially novel degraders.

Table 5.3: Putative bacterial toluene degraders identified based on 16S *rRNA* relative abundance response ratios between ^{13}C-toluene and ^{12}C-toluene labeled BD fraction samples.

Day	#	Highest classification	PICRUSt2 Predicted?	$^{13}C/^{12}C$ Labeled BD response ratio	Average relative abundance in ^{13}C-toluene labeled BD fractions (%)	Relative abundance in ^{13}C-toluene whole microcosm (%)[a]	Abundance in ^{13}C-toluene whole microcosm relative to Unamended control	Nearest relative according to NCBI (% identity)[b]	Known degrader (reference)	Accession number
13	1	Genus: *Cupriavidus*	Yes	30.52	11.02	15.62	2204.66	***Cupriavidus necator* (99.59%)**[3]	(Perez-Pantoja et al. 2008)	MN909835
	2	Genus: *Cupriavidus*	Yes	12.11	0.65	4.70	ND Unamended	***Cupriavidus necator* (100%)**[c]	(Perez-Pantoja et al. 2008)	MN909836
	3	Genus: *Rhodococcus*	Yes	1950.91	30.77	4.07	114.97	***Rhodococcus jostii* (100%)**	(Weathers et al. 2015)	MN909837
	4	Genus: *Rhodococcus*	Yes	393.14	15.18	3.70	ND Unamended	***Rhodococcus wratislaviensis* (100%)**[c]	(Auffret et al. 2009)	MN909838
	5	Genus: *Luteimonas*	Yes	3.99	1.21	3.65	ND Unamended	*Luteimonas vadosa* (99.59%)[c]		MN909839
	6	Family: *Burkholderiaceae*	Yes	6.13	1.13	2.49	9.49	*Acidovorax konjaci* (100%)		MN909840
	7	Genus: *Azoarcus*	Yes	23.20	4.73	2.40	153.70	***Azoarcus buckelii* (100%)**[c]	(Mechichi et al. 2002)	MN909841
	8	Family: *Cellulomonadaceae*	Yes	213.24	0.72	2.17	ND Unamended	*Actinotalea solisilvae* (99.59%)[c]		MN909842
	9	Family: *Burkholderiaceae*	Yes	6.20	0.38	1.80	ND Unamended	*Noviherbaspirillum canariense* (99.17%)[c]		MN909843
	10	Genus: *Ochrobactrum*	Yes	45.58	0.25	1.61	ND Unamended	*Ochrobactrum lupini* (100%)[c]		MN909844
	11	Genus: *Devosia*	Yes	4.64	0.30	1.16	16.39	*Devosia humi* (100%)[c]		MN909845
	12	Genus: *Herbaspirillum*	Yes	7.09	0.20	0.25	ND Unamended	***Herbaspirillum chlorophenolicum* (97.11%)**[c]	(Xu et al. 2016)	MN909846
	13	Family: *Bdellovibrionaceae* (OM27 clade)	**No**	5.21	0.19	0.22	ND Unamended	*Thioprofundum hispidum* (86.36%)		MN909847
	14	Family: *Burkholderiaceae*	Yes	7.62	0.13	0.19	ND Unamended	*Azohydromonas ureilytica* (99.17%)[c]		MN909848
	15	Family: *Intrasporangiaceae*	**No**	91.41	0.52	0.19	ND Unamended	*Janibacter indicus* (100%)[c]		MN909849
	16	Genus: Roseomonas	Yes	11.62	0.12	0.17	ND Unamended	*Roseomonas wooponensis* (99.17%)		MN909850
	17	Order: Gaiellales	Yes	4.65	0.03	0.04	ND Unamended	*Gaiella occulta* (92.98%)		MN909851
	18	Genus: *Cellulomonas*	Yes	3.90	0.01	ND	ND Unamended	*Cellulomonas humilata* (100%)[c]		MN909852

18	19 (2)	Genus: *Cupriavidus*	Yes	3.62	1.33	12.64	ND Unamended	**(see 2)**	MN909836
	20	Genus: *Pseudomonas*	Yes	45.08	19.25	8.75	ND Unamended	*Pseudomonas borbori* (99.59%)	MN909853
	21	Genus: *Pseudomonas*	Yes	47.53	11.58	5.00	ND Unamended	*Pseudomonas borbori* (98.76%)	MN909854
	22	Family: *Micrococcaceae*	Yes	3.28	3.22	3.00	0.65	*Pseudarthrobacter phenanthrenivorans* (100%)[c]	MN909855
	23	Family: *Burkholderiaceae*	Yes	4.69	1.21	0.52	ND Unamended	*Actimicrobium antarcticum* (99.59%)	MN909856
	24 (4)	Genus: *Rhodococcus*	Yes	150.27	3.57	0.51	ND Unamended	**(see 4)**	MN909838
	25 (3)	Genus: *Rhodococcus*	Yes	36.79	2.03	0.33	9.26	**(see 3)**	MN909837
	26	Genus: *Sphingomonas*	Yes	3.01	0.15	0.27	ND Unamended	*Sphingomonas sediminicola* (100%)	MN909857
	27	Genus: *Sphingomonas*	Yes	6.14	0.16	0.10	7.34	*Sphingomonas jatrophae* (97.11%)	MN909858
	28	Genus: *Rhodococcus*	Yes	3.93	0.06	0.06	ND Unamended	*Rhodococcus tukisamuensis* (100%)	MN909859
	29 (16)	Genus: *Roseomonas*	Yes	4.06	0.04	0.03	ND Unamended	(see 16)	MN909850
	30 (13)	Family: *Bdellovibrionaceae* (OM27 clade)	**No**	10.20	0.04	0.02	ND Unamended	(see 13)	MN909847
	31	Family: *Chitinophagaceae*	**No**	4.18	0.02	0.01	ND Unamended	*Rurimicrobium arvi* (93.39%)	MN909860
	32	Order: Sphingobacteriales (AKYH767)	**No**	6.26	0.01	ND	ND Unamended	*Rurimicrobium arvi* (84.84%)	MN909861
	33	Order: Thermomicrobiales (JG30-KF-CM45)	Yes	4.67	0.02	ND	--	*Thermorudis peleae* (88.07%)	MN909862

[1] ND = non-detect.
[2] Known degraders in bold.
[3] Additional matches at this % identity.

5.4.5 Putative fungal toluene degraders

Only two fungal species, *Mortierella alpina* and *Purpureocillium lilacinum*, exhibited ITS sequence response ratios indicative (>3-fold response) of inferred assimilation of ^{13}C-toluene (Table 5.4). One *Mortierella alpina* sequence variant demonstrated a 3-fold response ratio for day 18 but was below detection in the corresponding microcosm. Three other sequence variants classified to the same species were more strongly implicated as degraders; they were found in the labeled BD fractions and exhibited substantial enrichment (6.80-19.32%) in the corresponding toluene-amended microcosm (Table 5.4). Their response ratios were, however, below the 3-fold cut-off for identifying high-likelihood degraders or could not be calculated due to absence of the same sequence in the ^{12}C-toluene sample. Nevertheless, these results provide strong evidence, for the first time to the best of present knowledge, that *Mortierella alpina* is involved in biodegradation of toluene, and supports previous implication of this trait in the *Mortierella* genus. *Mortierella isabellina* was previously observed to transform toluene to benzyl alcohol through cytochrome P450-mediated hydroxylation (Holland et al. 1988), while another species, *Mortierella elongatus*, is believed also to have degradation capacity based on functional annotation of toluene degradation pathway genes in an *M. elongatus* genome assembly (Li et al. 2018). In the present study, *M. elongatus* was enriched in toluene-amended microcosms (15.88% of sequences on average) but not directly implicated in toluene degradation by DNA-SIP. On day 18, nearly half (49.07%) of the sequences on average in the labeled BD range belonged to strains of *Mortierella elongatus* (Figure 5.5B); interestingly, however, none of those sequences were detected in the day 18 ^{13}C-toluene-amended microcosm and thus did not belong to the *M. elongatus* strains in high relative abundance in that microcosm. Potential toluene-degrading activities of both *Mortierella alpina and M. elongatus* merit further investigation. *M. elongatus* is a known poplar endophyte capable of performing plant growth-promoting and defense-stimulating activities (Liao et al. 2019) and might therefore have particular promise as an inoculant in phytoremediation enhancement strategies (Jambon et al. 2018).

DNA-SIP provided evidence of toluene utilization also by the fungus *Purpureocillium lilacinum*. A sequence classified to this species was below detection in the Unamended control, but exhibited a 2.6-fold enrichment in the ^{13}C-toluene microcosm after 13 days of incubation and showed a strong response ratio in the labeled BD fractions compared to the ^{12}C-toluene control (Table 5.4). An entomopathogenous fungus, *P. lilacinus* has prior toluene remediation application having been employed in gas-phase toluene biofiltration due to its ability to utilize toluene as a singular carbon and energy source (Vigueras et al. 2008). This species might have a role in phytoremediation applications, but its potential as an enhancement inoculant might be limited by its clinical significance as an infrequent human pathogen (Khan et al. 2012) – though it is already used as a nematode biocontrol agent (Singh et al. 2013). *P. lilacinum* was recently identified as an endophytic species of the Pb-hyperaccumulating herb *Dysphania ambrosioides* at multiple heavy-metal contaminated sites (Parmar et al. 2018). Broader host-plant specificity is unclear, with conflicting evidence on the ability of *P. lilacinum* to establish as an endophyte of agricultural crops suggesting that this ability may be strain-dependent (Holland et al. 2003; Lan et al. 2017). Endophytic degraders can enhance phytoremediation, including through nutrient cycling, pathogen protection, plant growth promotion, and by reducing contaminant concentrations and thereby decreasing phytotoxicity (Correa-Garcia et al. 2018). The present study confirmed that toluene-degrading *P. lilacinum* can, at the least, inhabit rhizosphere soil in association with hybrid poplar roots under impacted field conditions. The ability of *P. lilacinum* to establish as an endophyte of poplar roots is unclear and merits further investigation.

Table 5.4: Putative fungal toluene degraders identified based on ITS *rRNA* response ratios between ^{13}C and ^{12}C control and/or observed enrichment after toluene amendment.

Day	#	Highest classification	$^{13}C/^{12}C$ Labeled BD response ratio	Average relative abundance in ^{13}C-toluene labeled BD fractions (%)	Relative abundance in ^{13}C-toluene whole microcosm (%)[a]	Abundance in ^{13}C-toluene whole microcosm relative to Unamended control	Accession number
13	1	*Purpureocillium lilacinum*[b]	32.25	0.32	2.60	ND Unamended	MN910273
18	2	*Mortierella alpina*	3.18	1.89	ND	--	MN910274
	3	*Mortierella alpina*	Infinite[c]	8.94	19.32	2.07	MN910275
	4	*Mortierella alpina*	1.81	8.67	18.46	ND Unamended	MN910276
	5	*Mortierella alpina*	1.70	3.06	6.80	ND Unamended	MN910277

[a] ND = non-detect.
[b] Organism is a known toluene degrader (Vigueras et al. 2008)
[c] Sequence not found in ^{12}C-toluene samples.

5.5 Conclusions

This exploratory study sought to identify toluene-degrading fungal and bacterial/archaeal organisms in rhizosphere soil obtained from a hybrid poplar phytoremediation system under loosely field-simulated conditions. Putative toluene-degrading bacteria were identified through DNA-SIP methods within the Actinobacteria and Proteobacteria phyla. The key degraders from these groups classified as *Cupriavidus*, *Rhodococcus*, *Luteimonas*, *Burkholderiaceae*, *Azoarcus*, *Cellulomonadaceae*, and *Pseudomonas* organisms. Sequences of putative degraders found in this study have been made available for future reference.

A novel element of this study, to the best of present knowledge, was the use of PICRUSt2 taxonomic functional modeling for the first time in a DNA-SIP study to characterize density-resolved fractions. Toluene degradation capacity was predicted for 16S *rRNA* gene sequences across the buoyant density gradient to provide additional lines of evidence for successful isolation of heavy isotope-labeled DNA. This technique adds to existing fraction characterization approaches, such as terminal restriction fragment length polymorphism (T-RFLP), DGGE, and direct stable isotope quantification (Dunford and Neufeld 2010; Wilhelm et al. 2014), and shows promise to support DNA-SIP studies assessing other metabolic pathways. Furthermore, PICRUSt2 was able to predict toluene

degradation capacity for the majority of sequences that were identified through DNA-SIP as coming from putative degraders. Sequences that were not found in the PICRUSt2-generated metagenomic libraries, classified within the *Bdellovibrionaceae* (OM27 clade), *Intrasporangiaceae*, or *Chitinophagaceae* families, or within the Sphingobacteriales (AKYH767) order, are potentially novel and those organisms merit further investigation.

Another objective of the present study was to identify toluene-degrading fungal organisms that might also have potential as poplar root inoculants to enhance aromatic hydrocarbon phytoremediation systems. Two fungal species, *Mortierella alpina* and *Purpureocillium lilacinum*, showed direct evidence of ^{13}C-toluene assimilation in this study. While *P. lilacinum* is a known degrader, to the best of present knowledge, this study has shown for the first time that *Mortierella alpina* is involved in biodegradation of toluene. *Mortierella elongatus*, a poplar endophyte, was also implicated as a potential toluene degrader in this study, though indirectly. Enrichment of this species after toluene amendment and observation of sequences from this organism in ^{13}C-labeled DNA samples gave evidence for some *M. elongatus* strains that might carry out toluene-degrading activities. This species has particular promise for enhancing poplar phytoremediation systems and should be studied further.

Overall, this work provides a reference of bacterial 16S and fungal ITS *rRNA* partial gene sequences for putative toluene degraders from a hybrid poplar toluene phytoremediation system. These sequences will aid further studies and assessments to positively identify toluene-degrading organisms for more effective site assessment and monitoring activities.

Acknowledgements:

This work was supported by BP, the University Consortium for Field-Focused Groundwater Contamination Research, and the Natural Sciences and Engineering Research Council of Canada (NSERC). Special thank-you to Alan Scheibner (BP

Canada) for ongoing support and contributions to the project. Thank-you to Dyanne Brewer, Armen Charchoglyan, James Longstaffe, Jonathan Gaiero, Micaela Tosi, Dasiel Obregón Alvarez, and Maria Gorecka for technical support.

6 Summary and Conclusions

This investigation provided field-level evidence of mechanisms of a hybrid poplar phytoremediation system active upon a toluene-impacted shallow fractured bedrock aquifer system. The combined approaches of CSIA and molecular analysis of toluene degradation functional genes and gene transcripts were used to determine if vadose zone biodegradation is an end-fate for toluene in the phytoremediation system. The enrichment and activity of functional toluene biodegradation marker genes, corroborated by toluene isotopic enrichment patterns, provided evidence that this is indeed occurring. Based on these functional gene profiles, resolved spatially and across plant ecological compartments, distinct degrader assemblages were found to be populating rhizosphere soil and root microbiomes across the site. Notably, selection of degraders in deep roots occurred in trees overlying more heavily-impacted groundwater at the site.

This relationship between root microbiome communities and toluene phytoextraction was explored by spatially resolving uptake patterns and phytoextraction rates. Passive *in planta* chemical monitoring confirmed that phytoextraction from the subsurface, and presumedly phytovolatilization, is an additional remediation mechanism active at this site. Phytoextraction rates during peak season were between 313 to 743 µg/day for trees with detectable concentrations of toluene. On a plot-wide scale, toluene concentrations in trees provided an approximation of subsurface impacts. On a tree-by-tree basis, they highlighted that multiple processes, including spatio-temporally variable contaminant mass and phase distribution, hydrology, and biodegradation capacity within root microbiomes are key determinants for how much toluene is phytoextracted to aboveground points of measurement. Further effort is required to resolve, in a mass-balance context, the contributing sources of toluene detected *in planta*, such that total mass reduction rates may be established and remedial time-frames evaluated. Though biodegradation rates were not directly established, ecological lines of evidence for activity were produced. A selective pressure of toluene access active upon root-colonizing

bacteria was apparent, particularly at locations with higher concentrations of toluene in groundwater. High toluene exposure produced ecologically distinct, more uniform communities with greater predicted and measured degradation capacity in roots overlying the highly impacted area of the site. Notable enrichment of *Streptomyces* observed in these roots points to its potential importance as a poplar colonizer in toluene phytoremediation. A similar selection effect was observed higher up in the transpiration stream.

Endophytic communities in the stem phyllosphere were assessed in trees across the site and found to be dynamic seasonally, increasing in diversity and richness in late summer compared to early spring. Distinct bacterial communities were found in the more-heavily impacted areas of the site, as with root microbiomes, and seasonal increases in their diversity and richness exceeded general site-wide effects. These communities resided in trees with the highest detected *in planta* toluene concentrations and demonstrated a greater predicted toluene degradation capacity. Thus, a selective effect and biodegradation response to toluene exposure in poplar transpiration streams was evidenced in poplar stem tissue. Toluene-selected microbiomes were enriched in late summer with *Rhodobacteraceae*, *Fibrobacter*, *Bacteroides ihuae*, *Rhizobiaceaee*, *Microbacteriaceae*, *Arcobacter butzleri*, *Pseudomonas*, *Pseudomonadaceae*, and *Elusimicrobium* phylotypes, whose taxonomic sequences have been made available for future ecological monitoring efforts.

A reliance on previously genetically- and taxonomically-characterized toluene-degrading organisms was implicit in the high throughout sequencing and taxonomic functional modeling techniques employed here to assess degraders *in situ*. To expand this knowledge and thereby further optimize these tools, stable isotope probing methods were applied to identify organisms implicated in toluene utilization in a targeted manner. Bacterial (*Cupriavidus*, *Rhodococcus*, *Luteimonas*, *Burkholderiaceae*, *Azoarcus*, *Cellulomonadaceae*, and *Pseudomonas*) and fungal (*Mortierella alpina*, *Purpureocillium*

lilacinum, and *Mortierella elongatus*) putative degraders were identified and their taxonomic sequences made available for future reference. Future efforts should be directed at culture-based isolation and characterization of these organisms to explore their potential as bioaugmentation agents to enhance phytoremediation systems.

Based on the results of this thesis, it can be concluded that phytoremediation processes – rhizodegradation and phytoextraction/phytovolatilization – are indeed viably active on toluene derived from an aged LNAPL source in fractured bedrock. This work provides further evidence, quantitatively, that aromatic hydrocarbon mass reduction through poplar phytoremediation systems occurs predominantly in the subsurface. The mass removal rates established in this thesis provide a reference point for future systems. However, it is recommended that, given in-tandem biodegradation that precedes uptake and that occurs throughout the transpiration stream, phytoextraction assessments should, where possible, be supported by characterization of microbe-mediated biodegradation in all relevant media. While assessment of plant transpiration influences on site hydrology were outside the scope of this thesis, transpiration rates generated in this work are being analyzed with complementary high temporal-resolution groundwater level data sets to establish the 3-D hydraulic zone of influence of the poplar stand (Guo et al. 2019).

This work has broadened the toolset of field-validated methodologies to study the presence and activity of microbial toluene-degrading communities in soil, roots, and above-ground plant tissue. It provides validation for emerging chemical sampling and metagenomic techniques that are fundamental for effective monitoring of contaminated sites and risk management programs. These results will aid in future efforts to enhance phytoremediation application and monitoring strategies, and to quantify processes to inform remedial time-frames.

REFERENCES

Alic S, Pedron J, Dreo T, Van Gijsegem F. 2019. Genomic characterisation of the new *Dickeya fangzhongdai* species regrouping plant pathogens and environmental isolates. BMC Genomics. 20(1):34. https://doi.org/10.1186/s12864-018-5332-3

Allen AS, Borchardt MA, Kieke BA, Dunfield KE, Parker BL. 2017. Virus occurrence in private and public wells in a fractured dolostone aquifer in Canada. Hydrogeology Journal. 25(4):1117-1136. https://doi.org/10.1007/s10040-017-1557-5

Apitz SE, Davis JW, Finkelstein K, Hohreiter DW, Hoke R, Jensen RH, Jersak J, Kirtay VJ, Mack EE, Magar VS et al. 2005. Assessing and managing contaminated sediments: Part II, evaluating risk and monitoring sediment remedy effectiveness. Integr Environ Asses. 1(1):E1-E14. https://doi.org/10.1897/IEAM_2004a-002e.1

Apprill A, McNally S, Parsons R, Weber L. 2015. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquatic Microbial Ecology. 75(2):129-137. https://doi.org/10.3354/ame01753

Arenz BE, Schlatter DC, Bradeen JM, Kinkel LL. 2015. Blocking primers reduce co-amplification of plant DNA when studying bacterial endophyte communities. J Microbiol Methods. 117:1-3. https://doi.org/10.1016/j.mimet.2015.07.003

Auffret M, Labbe D, Thouand G, Greer CW, Fayolle-Guichard F. 2009. Degradation of a mixture of hydrocarbons, gasoline, and diesel oil additives by *Rhodococcus aetherivorans* and *Rhodococcus wratislaviensis*. Appl Environ Microbiol. 75(24):7774-7782. https://doi.org/10.1128/AEM.01117-09

Balachandran C, Duraipandiyan V, Balakrishna K, Ignacimuthu S. 2012. Petroleum and polycyclic aromatic hydrocarbons (PAHs) degradation and naphthalene metabolism in *Streptomyces* sp. (ERI-CPDA-1) isolated from oil contaminated soil. Bioresour Technol. 112:83-90. https://doi.org/10.1016/j.biortech.2012.02.059

Baldwin BR, Biernacki A, Blair J, Purchase MP, Baker JM, Sublette K, Davis G, Ogles D. 2010. Monitoring gene expression to evaluate oxygen infusion at a gasoline-contaminated site. Environ Sci Technol. 44(17):6829-6834. https://doi.org/10.1021/es101356t

Baldwin BR, Nakatsu CH, Nies L. 2003. Detection and enumeration of aromatic oxygenase genes by multiplex and real-time PCR. Appl Environ Microbiol. 69(6):3350-3358. https://doi.org/10.1128/aem.69.6.3350-3358.2003

Baoune H, Ould El Hadj-Khelil A, Pucci G, Sineli P, Loucif L, Polti MA. 2018. Petroleum degradation by endophytic *Streptomyces* spp. isolated from plants grown in contaminated soil of southern Algeria. Ecotoxicol Environ Saf. 147:602-609. https://doi.org/10.1016/j.ecoenv.2017.09.013

Baquiran JP, Thater B, Songco K, Crowley DE. 2012. Characterization of culturable PAH and BTEX degrading bacteria from heavy oil of the Rancho La Brea Tar Pits.

Polycyclic Aromatic Compounds. 32(5):600-614. https://doi.org/10.1080/10406638.2011.651678

Barac T, Taghavi S, Borremans B, Provoost A, Oeyen L, Colpaert JV, Vangronsveld J, van der Lelie D. 2004. Engineered endophytic bacteria improve phytoremediation of water-soluble, volatile, organic pollutants. Nat Biotechnol. 22(5):583-588. https://doi.org/10.1038/nbt960

Barac T, Weyens N, Oeyen L, Taghavi S, van der Lelie D, Dubin D, Spliet M, Vangronsveld J. 2009. Field note: hydraulic containment of a BTEX plume using poplar trees. Int J Phytoremediation. 11(5):416-424. https://doi.org/10.1080/15226510802655880

Beckers B, Op De Beeck M, Weyens N, Boerjan W, Vangronsveld J. 2017. Structural variability and niche differentiation in the rhizosphere and endosphere bacterial microbiome of field-grown poplar trees. Microbiome. 5(1):25. https://doi.org/10.1186/s40168-017-0241-2

Bell TH, Yergeau E, Maynard C, Juck D, Whyte LG, Greer CW. 2013. Predictable bacterial composition and hydrocarbon degradation in Arctic soils following diesel and nutrient disturbance. ISME J. 7(6):1200-1210. https://doi.org/10.1038/ismej.2013.1

Beller HR, Kane SR, Legler TC, McKelvie JR, Lollar BS, Pearson F, Balser L, Mackay DM. 2008. Comparative assessments of benzene, toluene, and xylene natural attenuation by quantitative polymerase chain reaction analysis of a catabolic gene, signature metabolites, and compound-specific isotope analysis. Environ Sci Technol. 42(16):6065-6072. https://doi.org/10.1021/es8009666

BenIsrael M, Wanner P, Aravena R, Parker BL, Haack EA, Tsao DT, Dunfield KE. 2019. Toluene biodegradation in the vadose zone of a poplar phytoremediation system identified using metagenomics and toluene-specific stable carbon isotope analysis. Int J Phytoremediation. 21(1):60-69. https://doi.org/10.1080/15226514.2018.1523873

BenIsrael M, Wanner P, Fernandes J, Burken JG, Aravena R, Parker BL, Haack EA, Tsao DT, Dunfield KE. 2020. Quantification of toluene phytoextraction rates and microbial biodegradation functional profiles at a fractured bedrock phytoremediation site. Sci Total Environ. 707:135890. https://doi.org/10.1016/j.scitotenv.2019.135890

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F et al. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. https://doi.org/10.1038/s41587-019-0209-9

Bouchard D, Hohener P, Hunkeler D. 2008a. Carbon isotope fractionation during volatilization of petroleum hydrocarbons and diffusion across a porous medium: a

column experiment. Environ Sci Technol. 42(21):7801-7806. https://doi.org/10.1021/es800391s

Bouchard D, Hunkeler D, Gaganis P, Aravena R, Hohener P, Broholm MM, Kjeldsen P. 2008b. Carbon isotope fractionation during diffusion and biodegradation of petroleum hydrocarbons in the unsaturated zone: Field experiment at Vaerlose airbase, Denmark, and modeling. Environ Sci Technol. 42(2):596-601. https://doi.org/10.1021/es070718f

Bourdel G, Roy-Bolduc A, St-Arnaud M, Hijri M. 2016. Concentration of petroleum-hydrocarbon contamination shapes fungal endophytic community structure in plant roots. Front Microbiol. 7:685. https://doi.org/10.3389/fmicb.2016.00685

Braeckevelt M, Fischer A, Kastner M. 2012. Field applicability of Compound-Specific Isotope Analysis (CSIA) for characterization and quantification of in situ contaminant degradation in aquifers. Appl Microbiol Biotechnol. 94(6):1401-1421. https://doi.org/10.1007/s00253-012-4077-1

Brow CN, O'Brien Johnson R, Johnson RL, Simon HM. 2013. Assessment of anaerobic toluene biodegradation activity by *bssA* transcript/gene ratios. Appl Environ Microbiol. 79(17):5338-5344. https://doi.org/10.1128/AEM.01031-13

Buckley DH, Huangyutitham V, Hsu SF, Nelson TA. 2007. Stable isotope probing with ^{15}N achieved by disentangling the effects of genome G+C content and isotope enrichment on DNA density. Appl Environ Microbiol. 73(10):3189-3195. https://doi.org/10.1128/AEM.02609-06

Burken JG, Schnoor JL. 1998. Predictive relationships for uptake of organic contaminants by hybrid poplar trees. Environmental Science & Technology. 32(21):3379-3385. https://doi.org/10.1021/es9706817

Burken JG, Schnoor JL. 1999. Distribution and volatilization of organic compounds following uptake by hybrid poplar trees. Int J Phytoremediat. 1(2):139-151. https://doi.org/10.1080/15226519908500012

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL et al. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem. 55(4):611-622. https://doi.org/10.1373/clinchem.2008.112797

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods. 13(7):581-583. https://doi.org/10.1038/nmeth.3869

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M et al. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 6(8):1621-1624. https://doi.org/10.1038/ismej.2012.8

Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, Holland TA, Keseler IM, Kothari A, Kubo A et al. 2014. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. Nucleic Acids Res. 42(Database issue):D459-471. https://doi.org/10.1093/nar/gkt1103

CCME. 2004. Canadian soil quality guidelines for the protection of environmental and human health: Toluene (2004). Canadian environmental quality guidelines, 1999. Canadian Council of Ministers of the Environment.

Cho Y, Han K, Kim N, Park S, Kim Y. 2013. Estimating *in situ* biodegradation rates of petroleum hydrocarbons and microbial population dynamics by performing single-well push–pull tests in a fractured bedrock aquifer. Water, Air, & Soil Pollution. 224(2). https://doi.org/10.1007/s11270-012-1364-5

Christophersen M, Broholm MM, Mosbaek H, Karapanagioti HK, Burganos VN, Kjeldsen P. 2005. Transport of hydrocarbons from an emplaced fuel source experiment in the vadose zone at Airbase Vaerlose, Denmark. J Contam Hydrol. 81(1-4):1-33. https://doi.org/10.1016/j.jconhyd.2005.06.011

Compernolle T, Van Passel S, Weyens N, Vangronsveld J, Lebbe L, Thewys T. 2012. Groundwater remediation and the cost effectiveness of phytoremediation. Int J Phytoremediation. 14(9):861-877. https://doi.org/10.1080/15226514.2011.628879

Correa-Garcia S, Pande P, Seguin A, St-Arnaud M, Yergeau E. 2018. Rhizoremediation of petroleum hydrocarbons: a model system for plant microbiome manipulation. Microb Biotechnol. 11(5):819-832. https://doi.org/10.1111/1751-7915.13303

Crombie AT, Larke-Mejia NL, Emery H, Dawson R, Pratscher J, Murphy GP, McGenity TJ, Murrell JC. 2018. Poplar phyllosphere harbors disparate isoprene-degrading bacteria. Proc Natl Acad Sci U S A. 115(51):13081-13086. https://doi.org/10.1073/pnas.1812668115

Davee KW, Sanders DA. 2000. Petroleum hydrocarbon monitored natural attenuation: essential framework for remedial managers. Environmental Geosciences. 7(4):190-202. https://doi.org/10.1046/j.1526-0984.2000.74004.x

Dorn-In S, Bassitta R, Schwaiger K, Bauer J, Holzel CS. 2015. Specific amplification of bacterial DNA by optimized so-called universal bacterial primers in samples rich of plant DNA. J Microbiol Methods. 113:50-56. https://doi.org/10.1016/j.mimet.2015.04.001

Douglas G. 2019. PICRUSt2. [accessed 2019]. https://github.com/picrust/picrust2.

Douglas GM, Beiko RG, Langille MGI. 2018. Predicting the functional potential of the microbiome from marker genes using PICRUSt. In: Beiko RG, Hsiao W, Parkinson J, editors. Microbiome Analysis: Methods and Protocols, Methods in Molecular Biology. New York, New York, United States: Humana Press. p. 169-177.

Douglas GM, Maffei VJ, Zaneveld J, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI. 2019. PICRUSt2: An improved and extensible approach for metagenome inference. bioRxiv.672295. https://doi.org/10.1101/672295

Dunford EA, Neufeld JD. 2010. DNA stable-isotope probing (DNA-SIP). J Vis Exp. (42). https://doi.org/10.3791/2027

Durand A, Maillard F, Alvarez-Lopez V, Guinchard S, Bertheau C, Valot B, Blaudez D, Chalot M. 2018. Bacterial diversity associated with poplar trees grown on a Hg-contaminated site: community characterization and isolation of Hg-resistant plant growth-promoting bacteria. Sci Total Environ. 622-623:1165-1177. https://doi.org/10.1016/j.scitotenv.2017.12.069

El-Gendy AS, Svingos S, Brice D, Garretson JH, Schnoor J. 2009. Assessments of the efficacy of a long-term application of a phytoremediation system using hybrid poplar trees at former oil tank farm sites. Water Environ Res. 81(5):486-498. https://doi.org/10.2175/106143008x357011

Elora Research Station. 2017. Weather records for the Elora Research Station, Elora, Ontario [Canada]: Meteorological data 2017. In: Agricultural Forest Meteorology Group. Elora Research Station/Guelph Turfgrass Institute SoES, Ontario Agricultural College, University of Guelph, editor. V5 ed.: Scholars Portal Dataverse. https://doi.org/10.5683/SP/KYKL9M

Fahy A, Ball AS, Lethbridge G, Timmis KN, McGenity TJ. 2008. Isolation of alkali-tolerant benzene-degrading bacteria from a contaminated aquifer. Lett Appl Microbiol. 47(1):60-66. https://doi.org/10.1111/j.1472-765X.2008.02386.x

Fernandes J. 2017. Nature and extent of toluene contamination in a shallow dolostone aquifer using high resolution methods for assessing natural and anthropogenic influences [Masters]. [Guelph, Ontario, Canada]: University of Guelph.

Ferro A, Chard J, Kjelgren R, Chard B, Turner D, Montague T. 2001. Groundwater capture using hybrid poplar trees: evaluation of a system in Ogden, Utah. Int J Phytoremediat. 3(1):87-104. https://doi.org/10.1080/15226510108500051

Ferro A, Gefell, M., Kjelgren, R., Lipson, D. S., Zollinger, N., Jackson, S. 2003. Maintaining hydraulic control using deep rooted tree systems. In: Scheper T, editor. Advances in Biochemical Engineering/Biotechnology.

Ferro AM, Adham T, Berra B, Tsao D. 2013. Performance of deep-rooted phreatophytic trees at a site containing total petroleum hydrocarbons. Int J Phytoremediation. 15(3):232-244. https://doi.org/10.1080/15226514.2012.687195

Ficker M, Krastel K, Orlicky S, Edwards E. 1999. Molecular characterization of a toluene-degrading methanogenic consortium. Appl Environ Microbiol. 65(12):5576-5585.

Furukawa J, Abe Y, Mizuno H, Matsuki K, Sagawa K, Kojima M, Sakakibara H, Iwai H, Satoh S. 2011. Seasonal fluctuation of organic and inorganic components in xylem sap of *Populus nigra*. Plant Root. 5:56-62. https://doi.org/10.3117/plantroot.5.56

Gaiero JR, McCall CA, Thompson KA, Day NJ, Best AS, Dunfield KE. 2013. Inside the root microbiome: bacterial root endophytes and plant growth promotion. Am J Bot. 100(9):1738-1750. https://doi.org/10.3732/ajb.1200572

Garza DR, Dutilh BE. 2015. From cultured to uncultured genome sequences: metagenomics and modeling microbial ecosystems. Cell Mol Life Sci. 72(22):4287-4308. https://doi.org/10.1007/s00018-015-2004-1

Gatliff E, Linton PJ, Riddle DJ, Thomas PR. 2016. Phytoremediation of soil and groundwater: economic benefits over traditional methodologies. Bioremediation and Bioeconomy. Elsevier. p. 589-608.

Goli O, Górecki T, Mugammar H, Marchesi M, Aravena R. 2017. Evaluation of the suitability of the Waterloo Membrane Sampler for sample preconcentration before compound-specific isotope analysis. Environmental Technology & Innovation. 7:141-151. https://doi.org/10.1016/j.eti.2017.02.001

Gottel NR, Castro HF, Kerley M, Yang Z, Pelletier DA, Podar M, Karpinets T, Uberbacher E, Tuskan GA, Vilgalys R et al. 2011. Distinct microbial communities within the endosphere and rhizosphere of *Populus deltoides* roots across contrasting soil types. Appl Environ Microbiol. 77(17):5934-5944. https://doi.org/10.1128/AEM.05255-11

Granier A. 1985. Une nouvelle méthode pour la mesure du flux de sève brute dans le tronc des arbres. Annales des Sciences Forestières. 42(2):193-200. https://doi.org/10.1051/forest:19850204

Guo J, Kennel J, BenIsrael M, Wanner P, Parker B. 2019. Plant transpiration influences on groundwater in a shallow dolostone aquifer. Paper presented at: Geological Society of America Annual Meeting. Phoenix, Arizona, USA.

Habtewold J, Gordon R, Sokolov V, VanderZaag A, Wagner-Riddle C, Dunfield K. 2018. Targeting bacteria and methanogens to understand the role of residual slurry as an inoculant in stored liquid dairy manure. Appl Environ Microbiol. 84(7). https://doi.org/10.1128/AEM.02830-17

Hardisty PE, Roher J, Dottridge J. 2004. LNAPL behavior in fractured rock: Implications for characterization and remediation. Paper presented at: US EPA/NGWA Fractured Rock Conference: State of the Science and Measuring Success in Remediation. Portland, Maine, USA.

Hirsch PR, Mauchline TH, Clark IM. 2010. Culture-independent molecular techniques for soil microbial ecology. Soil Biology and Biochemistry. 42(6):878-887. https://doi.org/10.1016/j.soilbio.2010.02.019

Hohener P, Dakhel N, Christophersen M, Broholm M, Kjeldsen P. 2006. Biodegradation of hydrocarbons vapors: comparison of laboratory studies and field investigations in the vadose zone at the emplaced fuel source experiment, Airbase Vaerlose,

Denmark. J Contam Hydrol. 88(3-4):337-358. https://doi.org/10.1016/j.jconhyd.2006.07.007

Holland HL, Brown FM, Munoz B, Ninniss RW. 1988. Side chain hydroxylation of aromatic hydrocarbons by fungi. Part 2. Isotope effects and mechanism. Journal of the Chemical Society, Perkin Transactions 2. (8):1557. https://doi.org/10.1039/p29880001557

Holland RJ, Williams KL, Nevalainen KMH. 2003. *Paecilomyces lilacinus* strain Bioact251 is not a plant endophyte. Australasian Plant Pathology. 32(4):473. https://doi.org/10.1071/ap03046

Holliger C, Gaspard S, Glod G, Heijman C, Schumacher W, Schwarzenbach RP, Vazquez F. 1997. Contaminated environments in the subsurface and bioremediation: organic contaminants. FEMS Microbiol Rev. 20(3-4):517-523. https://doi.org/10.1111/j.1574-6976.1997.tb00334.x

Hubbert KR, Beyers JL, Graham RC. 2001. Roles of weathered bedrock and soil in seasonal water relations of *Pinus Jeffreyi* and *Arctostaphylos patula*. Can J Forest Res. 31(11):1947-1957. https://doi.org/10.1139/cjfr-31-11-1947

Illman WA, Alvarez PJ. 2009. Performance assessment of bioremediation and natural attenuation. Crit Rev Env Sci Tec. 39(4):209-270. https://doi.org/10.1080/10643380701413385

ITRC. 2009a. Evaluating natural source zone depletion at sites with LNAPL. Washington, D.C.: Interstate Technology & Regulatory Council, LNAPLs Team.

ITRC. 2009b. Phytotechnology technical and regulatory guidance and decision trees, revised. Washington, D.C.: Interstate Technology & Regulatory Council, Phytotechnologies Team, Tech Reg Update.

Iwai S, Johnson TA, Chai B, Hashsham SA, Tiedje JM. 2011. Comparison of the specificities and efficacies of primers for aromatic dioxygenase gene analysis of environmental samples. Appl Environ Microbiol. 77(11):3551-3557. https://doi.org/10.1128/AEM.00331-11

Jambon I, Thijs S, Weyens N, Vangronsveld J. 2018. Harnessing plant-bacteria-fungi interactions to improve plant growth and degradation of organic pollutants. Journal of Plant Interactions. 13(1):119-130. https://doi.org/10.1080/17429145.2018.1441450

Jobbágy EG, Jackson RB. 2001. The distribution of soil nutrients with depth: global patterns and the imprint of plants. Biogeochemistry. 53(1):51-77. https://doi.org/10.1023/A:1010760720215

Jordahl JL, Foster L, Schnoor JL, Alvarez PJJ. 1997. Effect of hybrid poplar trees on microbial populations important to hazardous waste bioremediation. Environ Toxicol Chem. 16(6):1318-1321. https://doi.org/10.1002/etc.5620160630

Ju HJ, Hill NS, Abbott T, Ingram KT. 2006. Temperature influences on endophyte growth in tall fescue. Crop Science. 46(1):404. https://doi.org/10.2135/cropsci2005.0282

Julio ADL, de Cassia Mourao Silva U, Medeiros JD, Morais DK, Dos Santos VL. 2019. Metataxonomic analyses reveal differences in aquifer bacterial community as a function of creosote contamination and its potential for contaminant remediation. Sci Rep. 9(1):11731. https://doi.org/10.1038/s41598-019-47921-y

Kandel S, Joubert P, Doty S. 2017. Bacterial endophyte colonization and distribution within plants. Microorganisms. 5(4). https://doi.org/10.3390/microorganisms5040077

Kao CM, Chen CS, Tsa FY, Yang KH, Chien CC, Liang SH, Yang CA, Chen SC. 2010. Application of real-time PCR, DGGE fingerprinting, and culture-based method to evaluate the effectiveness of intrinsic bioremediation on the control of petroleum-hydrocarbon plume. J Hazard Mater. 178(1-3):409-416. https://doi.org/10.1016/j.jhazmat.2010.01.096

Karr JR, Chu EW. 2008. Biological monitoring: essential foundation for ecological risk assessment. Human and Ecological Risk Assessment: An International Journal. 3(6):993-1004. https://doi.org/10.1080/10807039709383742

Key KC, Sublette KL, Johannes TW, Ogles D, Baldwin B, Biernacki A. 2014. Assessing BTEX biodegradation potential at a refinery using molecular biological tools. Ground Water Monit R. 34(1):35-48. https://doi.org/10.1111/gwmr.12037

Khan Z, Ahmad S, Al-Ghimlas F, Al-Mutairi S, Joseph L, Chandy R, Sutton DA, Guarro J. 2012. *Purpureocillium lilacinum* as a cause of cavitary pulmonary disease: a new clinical presentation and observations on atypical morphologic characteristics of the isolate. J Clin Microbiol. 50(5):1800-1804. https://doi.org/10.1128/JCM.00150-12

Kim J, Corapcioglu MY. 2003. Modeling dissolution and volatilization of LNAPL sources migrating on the groundwater table. J Contam Hydrol. 65(1-2):137-158. https://doi.org/10.1016/S0169-7722(02)00105-5

Kostner B, Granier A, Cermak J. 1998. Sapflow measurements in forest stands: methods and uncertainties. Annales Des Sciences Forestieres. 55(1-2):13-27. https://doi.org/10.1051/forest:19980102

Krzmarzick MJ, Taylor DK, Fu X, McCutchan AL. 2018. Diversity and niche of archaea in bioremediation. Archaea. 2018:3194108. https://doi.org/10.1155/2018/3194108

Lan X, Zhang J, Zong Z, Ma Q, Wang Y. 2017. Evaluation of the biocontrol potential of *Purpureocillium lilacinum* QLP12 against *Verticillium dahliae* in eggplant. Biomed Res Int. 2017:4101357. https://doi.org/10.1155/2017/4101357

Landmeyer JE, Effinger TN. 2016. Effect of phytoremediation on concentrations of benzene, toluene, naphthalene, and dissolved oxygen in groundwater at a former

manufactured gas plant site, Charleston, South Carolina, USA, 1998–2014. Environmental Earth Sciences. 75(7). https://doi.org/10.1007/s12665-016-5408-9

Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Vega Thurber RL, Knight R et al. 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol. 31(9):814-821. https://doi.org/10.1038/nbt.2676

Langille MGI. 2018. Exploring linkages between taxonomic and functional profiles of the human microbiome. Msystems. 3(2). https://doi.org/10.1128/mSystems.00163-17

Larentis M, Hoermann K, Lueders T. 2013. Fine-scale degrader community profiling over an aerobic/anaerobic redox gradient in a toluene-contaminated aquifer. Environ Microbiol Rep. 5(2):225-234. https://doi.org/10.1111/1758-2229.12004

Li A, He W. 2019. Molecular aspects of an emerging poplar canker caused by *Lonsdalea populi*. Frontiers in Microbiology. 10. https://doi.org/10.3389/fmicb.2019.02496

Li F, Chen L, Redmile-Gordon M, Zhang J, Zhang C, Ning Q, Li W. 2018. *Mortierella elongata*'s roles in organic agriculture and crop growth promotion in a mineral soil. Land Degradation & Development. 29(6):1642-1651. https://doi.org/10.1002/ldr.2965

Li Y, Fang W, Xue H, Liang WX, Wang LF, Tian GZ, Wang XZ, Lin CL, Li X, Piao CG. 2015. *Brenneria populi* sp. nov., isolated from symptomatic bark of *Populus*×*euramericana* canker. Int J Syst Evol Microbiol. 65(Pt 2):432-437. https://doi.org/10.1099/ijs.0.066068-0

Liao HL, Bonito G, Rojas JA, Hameed K, Wu S, Schadt CW, Labbe J, Tuskan GA, Martin F, Grigoriev IV et al. 2019. Fungal endophytes of *Populus trichocarpa* alter host phenotype, gene expression, and rhizobiome composition. Mol Plant Microbe Interact. 32(7):853-864. https://doi.org/10.1094/MPMI-05-18-0133-R

Lima-Morales D, Jauregui R, Camarinha-Silva A, Geffers R, Pieper DH, Vilchez-Vargas R. 2016. Linking microbial community and catabolic gene structures during the adaptation of three contaminated soils under continuous long-term pollutant stress. Appl Environ Microbiol. 82(7):2227-2237. https://doi.org/10.1128/AEM.03482-15

Limmer M, Burken J. 2016. Phytovolatilization of organic contaminants. Environ Sci Technol. 50(13):6632-6643. https://doi.org/10.1021/acs.est.5b04113

Limmer MA, Balouet JC, Karg F, Vroblesky DA, Burken JG. 2011. Phytoscreening for chlorinated solvents using rapid in vitro SPME sampling: application to urban plume in Verl, Germany. Environ Sci Technol. 45(19):8276-8282. https://doi.org/10.1021/es201704v

Limmer MA, Burken JG. 2015. Phytoscreening with SPME: variability analysis. Int J Phytoremediation. 17(11):1115-1122. https://doi.org/10.1080/15226514.2015.1045127

Limmer MA, Holmes AJ, Burken JG. 2014a. Phytomonitoring of chlorinated ethenes in trees: a four-year study of seasonal chemodynamics *in planta*. Environ Sci Technol. 48(18):10634-10640. https://doi.org/10.1021/es502680p

Limmer MA, Martin GD, Watson CJ, Martinez C, Burken JG. 2014b. Phytoscreening: a comparison of *in planta* portable GC-MS and in vitro analyses. Ground Water Monit R. 34(1):49-56. https://doi.org/10.1111/gwmr.12039

Limmer MA, Shetty MK, Markus S, Kroeker R, Parker BL, Martinez C, Burken JG. 2013. Directional phytoscreening: contaminant gradients in trees for plume delineation. Environ Sci Technol. 47(16):9069-9076. https://doi.org/10.1021/es400437q

Limmer MA, West DM, Mu RP, Shi HL, Whitlock K, Burken JG. 2015. Phytoscreening for perchlorate: rapid analysis of tree sap. Environ Sci-Wat Res. 1(2):138-145. https://doi.org/10.1039/c4ew00103f

Limmer MA, Wilson J, Westenberg D, Lee A, Siegman M, Burken JG. 2018. Phytoremediation removal rates of benzene, toluene, and chlorobenzene. Int J Phytoremediation. 20(7):666-674. https://doi.org/10.1080/15226514.2017.1413330

Link CM, Thevathasan NV, Gordon AM, Isaac ME. 2015. Determining tree water acquisition zones with stable isotopes in a temperate tree-based intercropping system. Agroforestry Systems. 89(4):611-620. https://doi.org/10.1007/s10457-015-9795-9

Lovley DR. 2003. Cleaning up with genomics: applying molecular biology to bioremediation. Nat Rev Microbiol. 1(1):35-44. https://doi.org/10.1038/nrmicro731

Lundblad M, Lindroth A. 2002. Stand transpiration and sapflow density in relation to weather, soil moisture and stand characteristics. Basic and Applied Ecology. 3(3):229-243. https://doi.org/10.1078/1439-1791-00099

Lunsmann V, Kappelmeyer U, Benndorf R, Martinez-Lavanchy PM, Taubert A, Adrian L, Duarte M, Pieper DH, von Bergen M, Muller JA et al. 2016a. In situ protein-SIP highlights *Burkholderiaceae* as key players degrading toluene by para ring hydroxylation in a constructed wetland model. Environ Microbiol. 18(4):1176-1186. https://doi.org/10.1111/1462-2920.13133

Lunsmann V, Kappelmeyer U, Taubert A, Nijenhuis I, von Bergen M, Heipieper HJ, Muller JA, Jehmlich N. 2016b. Aerobic toluene degraders in the rhizosphere of a constructed wetland model show diurnal polyhydroxyalkanoate metabolism. Appl Environ Microbiol. 82(14):4126-4132. https://doi.org/10.1128/AEM.00493-16

Ma B, Hibbing ME, Kim HS, Reedy RM, Yedidia I, Breuer J, Breuer J, Glasner JD, Perna NT, Kelman A et al. 2007. Host range and molecular phylogenies of the soft rot Enterobacterial genera *Pectobacterium* and *Dickeya*. Phytopathology. 97(9):1150-1163. https://doi.org/10.1094/PHYTO-97-9-1150

Matthews DW, Massmann J, Strand SE. 2003. Influence of aquifer properties on phytoremediation effectiveness. Ground Water. 41(1):41-47. https://doi.org/10.1111/j.1745-6584.2003.tb02566.x

Mechichi T, Stackebrandt E, Gad'on N, Fuchs G. 2002. Phylogenetic and metabolic diversity of bacteria degrading aromatic compounds under denitrifying conditions, and description of *Thauera phenylacetica* sp. nov., *Thauera aminoaromatica* sp. nov., and *Azoarcus buckelii* sp. nov. Arch Microbiol. 178(1):26-35. https://doi.org/10.1007/s00203-002-0422-6

MOECP. 2011. Soil, ground water and sediment standards for use under Part XV.1 of the Environmental Protection Act. In: Ministry of the Environment C, and Parks, editor. Ontario, Canada: Ontario Ministry of the Environment.

Mohamed EF, Awad G, Andriantsiferana C, El-Diwany AI. 2016. Biofiltration technology for the removal of toluene from polluted air using *Streptomyces griseus*. Environ Technol. 37(10):1197-1207. https://doi.org/10.1080/09593330.2015.1107623

Moore FP, Barac T, Borremans B, Oeyen L, Vangronsveld J, van der Lelie D, Campbell CD, Moore ER. 2006. Endophytic bacterial diversity in poplar trees growing on a BTEX-contaminated site: the characterisation of isolates with potential to enhance phytoremediation. Syst Appl Microbiol. 29(7):539-556. https://doi.org/10.1016/j.syapm.2005.11.012

Morales P, Caceres M, Scott F, Diaz-Robles L, Aroca G, Vergara-Fernandez A. 2017. Biodegradation of benzo[alpha]pyrene, toluene, and formaldehyde from the gas phase by a consortium of *Rhodococcus erythropolis* and *Fusarium solani*. Appl Microbiol Biotechnol. 101(17):6765-6777. https://doi.org/10.1007/s00253-017-8400-8

Mukherjee A, Chettri B, Langpoklakpam JS, Basak P, Prasad A, Mukherjee AK, Bhattacharyya M, Singh AK, Chattopadhyay D. 2017. Bioinformatic approaches including predictive metagenomic profiling reveal characteristics of bacterial response to petroleum hydrocarbon contamination in diverse environments. Sci Rep. 7(1):1108. https://doi.org/10.1038/s41598-017-01126-3

Nebe J, Baldwin BR, Kassab RL, Nies L, Nakatsu CH. 2009. Quantification of aromatic oxygenase genes to evaluate enhanced bioremediation by oxygen releasing materials at a gasoline-contaminated site. Environ Sci Technol. 43(6):2029-2034. https://doi.org/10.1021/es900146f

Neufeld JD, Vohra J, Dumont MG, Lueders T, Manefield M, Friedrich MW, Murrell JC. 2007. DNA stable-isotope probing. Nat Protoc. 2(4):860-866. https://doi.org/10.1038/nprot.2007.109

Nolvak H, Sildvee T, Knipsalu M, Truu J. 2012. Application of microbial community profiling and functional gene detection for assessment of natural attenuation of petroleum hydrocarbons in boreal subsurface. Boreal Environ Res. 17(2):113-127. https://doi.org/10138/229837

OECD. 2012. SIDS Initial Assessment Profiles agreed in the course of the OECD HPV Chemicals Programme from 1993-2011. Organisation for Economic Co-operation and Development. No. ENV/JM/MONO(2012)4/PART3.

Ortega-Gonzalez DK, Zaragoza D, Aguirre-Garrido J, Ramirez-Saad H, Hernandez-Rodriguez C, Jan-Roblero J. 2013. Degradation of benzene, toluene, and xylene isomers by a bacterial consortium obtained from rhizosphere soil of *Cyperus* sp. grown in a petroleum-contaminated area. Folia Microbiol (Praha). 58(6):569-577. https://doi.org/10.1007/s12223-013-0248-4

Palau J, Marchesi M, Chambon JC, Aravena R, Canals A, Binning PJ, Bjerg PL, Otero N, Soler A. 2014. Multi-isotope (carbon and chlorine) analysis for fingerprinting and site characterization at a fractured bedrock aquifer contaminated by chlorinated ethenes. Sci Total Environ. 475:61-70. https://doi.org/10.1016/j.scitotenv.2013.12.059

Palleroni NJ, Pieper DH, Moore ERB. 2010. Microbiology of hydrocarbon-degrading *Pseudomonas*. In: Timmis KN, editor. Handbook of Hydrocarbon and Lipid Microbiology. Berlin Heidelberg: Springer-Verlag. p. 1787-1798.

Parada AE, Needham DM, Fuhrman JA. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ Microbiol. 18(5):1403-1414. https://doi.org/10.1111/1462-2920.13023

Parales RE. 2010. Hydrocarbon degradation by betaproteobacteria. In: Timmis KN, editor. Handbook of hydrocarbon and lipid microbiology. Berlin, Heidelberg: Springer. p. 1715-1724.

Parales RE, Parales JV, Pelletier DA, Ditty JL. 2008. Diversity of microbial toluene degradation pathways. Adv Appl Microbiol. 64:1-73. https://doi.org/10.1016/S0065-2164(08)00401-2

Parker BL, Cherry JA, Chapman SW. 2012. Discrete fracture network approach for studying contamination in fractured rock. AQUA mundi. 3(2):101-116. https://doi.org/10.4409/Am-052-12-0046

Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol. 36(10):996-1004. https://doi.org/10.1038/nbt.4229

Parks DH, Tyson GW, Hugenholtz P, Beiko RG. 2014. STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics. 30(21):3123-3124. https://doi.org/10.1093/bioinformatics/btu494

Parmar S, Li Q, Wu Y, Li X, Yan J, Sharma VK, Wei Y, Li H. 2018. Endophytic fungal community of *Dysphania ambrosioides* from two heavy metal-contaminated sites:

evaluated by culture-dependent and culture-independent approaches. Microb Biotechnol. 11(6):1170-1183. https://doi.org/10.1111/1751-7915.13308

Perez-Pantoja D, De la Iglesia R, Pieper DH, Gonzalez B. 2008. Metabolic reconstruction of aromatic compounds degradation from the genome of the amazing pollutant-degrading bacterium *Cupriavidus necator* JMP134. FEMS Microbiol Rev. 32(5):736-794. https://doi.org/10.1111/j.1574-6976.2008.00122.x

Piterina AV, Pembroke JT. 2013. Use of PCR-DGGE based molecular methods to analyse microbial community diversity and stability during the thermophilic stages of an ATAD wastewater sludge treatment process as an aid to performance monitoring. ISRN Biotechnol. 2013:162645. https://doi.org/10.5402/2013/162645

Prenafeta-Boldu FX, Summerbell R, Sybren de Hoog G. 2006. Fungi growing on aromatic hydrocarbons: biotechnology's unexpected encounter with biohazard? FEMS Microbiol Rev. 30(1):109-130. https://doi.org/10.1111/j.1574-6976.2005.00007.x

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41(Database issue):D590-596. https://doi.org/10.1093/nar/gks1219

Rakoczy J, Remy B, Vogt C, Richnow HH. 2011. A bench-scale constructed wetland as a model to characterize benzene biodegradation processes in freshwater wetlands. Environ Sci Technol. 45(23):10036-10044. https://doi.org/10.1021/es2026196

Reardon CL, Strauss SL, Mazzola M. 2013. Changes in available nitrogen and nematode abundance in response to *Brassica* seed meal amendment of orchard soil. Soil Biol Biochem. 57:22-29. https://doi.org/10.1016/j.soilbio.2012.10.011

Rivers AR, Weber KC, Gardner TG, Liu S, Armstrong SD. 2018. ITSxpress: software to rapidly trim internally transcribed spacer sequences with quality scores for marker gene analysis. F1000Res. 7:1418. https://doi.org/10.12688/f1000research.15704.1

Roebuck AJ. 2018. Assessment of toluene biodegradation activity in groundwater from a shallow bedrock aquifer underlying a phytoremediation system [Masters]. [Guelph, Ontario, Canada]: University of Guelph.

Rogers E, Zalesny R, Hallett R, Headlee W, Wiese A. 2019. Relationships among root–shoot ratio, early growth, and health of hybrid poplar and willow clones grown in different landfill soils. Forests. 10(1):49. https://doi.org/10.3390/f10010049

Roy A, Sar P, Sarkar J, Dutta A, Sarkar P, Gupta A, Mohapatra B, Pal S, Kazy SK. 2018. Petroleum hydrocarbon rich oil refinery sludge of North-East India harbours anaerobic, fermentative, sulfate-reducing, syntrophic and methanogenic microbial populations. BMC Microbiol. 18(1):151. https://doi.org/10.1186/s12866-018-1275-8

Saleh-Lakha S, Miller M, Campbell RG, Schneider K, Elahimanesh P, Hart MM, Trevors JT. 2005. Microbial gene expression in soil: methods, applications and challenges. J Microbiol Methods. 63(1):1-19. https://doi.org/10.1016/j.mimet.2005.03.007

Sanemasa I, Araki M, Deguchi T, Nagai H. 1982. Solubility measurements of benzene and the alkylbenzenes in water by making use of solute vapor. Bulletin of the Chemical Society of Japan. 55(4):1054-1062. https://doi.org/10.1246/bcsj.55.1054

Schumacher JG, Struckhoff GC, Burken JG. 2004. Assessment of subsurface chlorinated solvent contamination using tree cores at the Front Street site and a former dry cleaning facility at the Riverfront Superfund Site, New Haven, Missouri, 1999–2003.

Schwitzguébel J-P, van der Lelie D, Baker A, Glass DJ, Vangronsveld J. 2002. Phytoremediation: European and American trends successes, obstacles and needs. Journal of Soils and Sediments. 2(2):91-99. https://doi.org/10.1007/bf02987877

Seethapathy S, Gorecki T. 2011. Polydimethylsiloxane-based permeation passive air sampler. Part I: Calibration constants and their relation to retention indices of the analytes. J Chromatogr A. 1218(1):143-155. https://doi.org/10.1016/j.chroma.2010.11.003

Shakya M, Gottel N, Castro H, Yang ZK, Gunter L, Labbe J, Muchero W, Bonito G, Vilgalys R, Tuskan G et al. 2013. A multifactor analysis of fungal and bacterial community structure in the root microbiome of mature *Populus deltoides* trees. PLoS One. 8(10):e76382. https://doi.org/10.1371/journal.pone.0076382

Shen SY, Fulthorpe R. 2015. Seasonal variation of bacterial endophytes in urban trees. Front Microbiol. 6:427. https://doi.org/10.3389/fmicb.2015.00427

Shetty MK, Limmer MA, Waltermire K, Morrison GC, Burken JG. 2014. *In planta* passive sampling devices for assessing subsurface chlorinated solvents. Chemosphere. 104:149-154. https://doi.org/10.1016/j.chemosphere.2013.10.084

Shinoda Y, Sakai Y, Uenishi H, Uchihashi Y, Hiraishi A, Yukawa H, Yurimoto H, Kato N. 2004. Aerobic and anaerobic toluene degradation by a newly isolated denitrifying bacterium, *Thauera* sp. Strain DNT-1. Applied and Environmental Microbiology. 70(3):1385-1392. https://doi.org/10.1128/aem.70.3.1385-1392.2004

Siciliano SD, Germida JJ, Banks K, Greer CW. 2002. Changes in microbial community composition and function during a polyaromatic hydrocarbon phytoremediation field trial. Applied and Environmental Microbiology. 69(1):483-489. https://doi.org/10.1128/aem.69.1.483-489.2003

Singh S, Pandey RK, Goswami BK. 2013. Bio-control activity of *Purpureocillium lilacinumstrains* in managing root-knot disease of tomato caused by *Meloidogyne incognita*. Biocontrol Science and Technology. 23(12):1469-1489. https://doi.org/10.1080/09583157.2013.840770

Smith CJ, Osborn AM. 2009. Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. FEMS Microbiol Ecol. 67(1):6-20. https://doi.org/10.1111/j.1574-6941.2008.00629.x

Spini G, Spina F, Poli A, Blieux AL, Regnier T, Gramellini C, Varese GC, Puglisi E. 2018. Molecular and microbiological insights on the enrichment procedures for the isolation of petroleum degrading bacteria and fungi. Front Microbiol. 9:2543. https://doi.org/10.3389/fmicb.2018.02543

Stefani FO, Bell TH, Marchand C, de la Providencia IE, El Yassimi A, St-Arnaud M, Hijri M. 2015. Culture-dependent and -independent methods capture different microbial community fractions in hydrocarbon-contaminated soils. PLoS One. 10(6):e0128272. https://doi.org/10.1371/journal.pone.0128272

Taghavi S, Barac T, Greenberg B, Borremans B, Vangronsveld J, van der Lelie D. 2005. Horizontal gene transfer to endogenous endophytic bacteria from poplar improves phytoremediation of toluene. Appl Environ Microbiol. 71(12):8500-8505. https://doi.org/10.1128/AEM.71.12.8500-8505.2005

Tardif S, Yergeau E, Tremblay J, Legendre P, Whyte LG, Greer CW. 2016. The willow microbiome is influenced by soil petroleum-hydrocarbon concentration with plant compartment-specific effects. Front Microbiol. 7:1363. https://doi.org/10.3389/fmicb.2016.01363

Techtmann SM, Hazen TC. 2016. Metagenomic applications in environmental monitoring and bioremediation. J Ind Microbiol Biotechnol. 43(10):1345-1354. https://doi.org/10.1007/s10295-016-1809-8

Thierrin J, Davis GB, Barber C, Patterson BM, Pribac F, Power TR, Lambert M. 2009. Natural degradation rates of BTEX compounds and naphthalene in a sulphate reducing groundwater environment. Hydrological Sciences Journal. 38(4):309-322. https://doi.org/10.1080/02626669309492677

Thijs S, Sillen W, Rineau F, Weyens N, Vangronsveld J. 2016. Towards an enhanced understanding of plant-microbiome interactions to improve phytoremediation: engineering the metaorganism. Front Microbiol. 7:341. https://doi.org/10.3389/fmicb.2016.00341

Thijs S, Sillen W, Weyens N, Vangronsveld J. 2017. Phytoremediation: state-of-the-art and a key role for the plant microbiome in future trends and research prospects. Int J Phytoremediation. 19(1):23-38. https://doi.org/10.1080/15226514.2016.1216076

Timm CM, Campbell AG, Utturkar SM, Jun SR, Parales RE, Tan WA, Robeson MS, Lu TYS, Jawdy S, Brown SD et al. 2015. Metabolic functions of *Pseudomonas fluorescens* strains from *Populus deltoides* depend on rhizosphere or endosphere isolation compartment. Frontiers in Microbiology. 6. https://doi.org/10.3389/fmicb.2015.01118

Ulrich K, Ulrich A, Ewald D. 2008. Diversity of endophytic bacterial communities in poplar grown under field conditions. FEMS Microbiol Ecol. 63(2):169-180. https://doi.org/10.1111/j.1574-6941.2007.00419.x

UNITE Community. 2019. UNITE QIIME release for Fungi Version 18.11.2018.: UNITE Community.

van der Lelie D, Taghavi S, Monchy S, Schwender J, Miller L, Ferrieri R, Rogers A, Wu X, Zhu W, Weyens N et al. 2009. Poplar and its bacterial endophytes: coexistence and harmony. Critical Reviews in Plant Sciences. 28(5):346-358. https://doi.org/10.1080/07352680903241204

Vigueras G, Shirai K, Martins de Souza D, Franco TT, Fleuri LF, Revah S. 2008. Toluene gas phase biofiltration by *Paecilomyces lilacinus* and isolation and identification of a hydrophobin protein produced thereof. Appl Microbiol Biotechnol. 80(1):147-154. https://doi.org/10.1007/s00253-008-1490-6

von Netzer F, Pilloni G, Kleindienst S, Kruger M, Knittel K, Grundger F, Lueders T. 2013. Enhanced gene detection assays for fumarate-adding enzymes allow uncovering of anaerobic hydrocarbon degraders in terrestrial and marine systems. Appl Environ Microbiol. 79(2):543-552. https://doi.org/10.1128/AEM.02362-12

Vurukonda S, Giovanardi D, Stefani E. 2018. Plant growth promoting and biocontrol activity of *Streptomyces* spp. as endophytes. Int J Mol Sci. 19(4). https://doi.org/10.3390/ijms19040952

Wang Y, Zhang W, Ding C, Zhang B, Huang Q, Huang R, Su X. 2019. Endophytic communities of transgenic poplar were determined by the environment and niche rather than by transgenic events. Front Microbiol. 10:588. https://doi.org/10.3389/fmicb.2019.00588

Wanner P, Aravena R, Fernandes J, BenIsrael M, Haack EA, Tsao DT, Dunfield KE, Parker BL. 2019. Assessing toluene biodegradation under temporally varying redox conditions in a fractured bedrock aquifer using stable isotope methods. Water Research.114986. https://doi.org/10.1016/j.watres.2019.114986

Wanner P, Hunkeler D. 2015. Carbon and chlorine isotopologue fractionation of chlorinated hydrocarbons during diffusion in water and low permeability sediments. Geochim Cosmochim Acta. 157:198-212. https://doi.org/10.1016/j.gca.2015.02.034

Weathers TS, Higgins CP, Sharp JO. 2015. Enhanced biofilm production by a toluene-degrading *Rhodococcus* observed after exposure to perfluoroalkyl acids. Environ Sci Technol. 49(9):5458-5466. https://doi.org/10.1021/es5060034

Weishaar JA, Tsao D, Burken JG. 2009. Phytoremediation of BTEX hydrocarbons: potential impacts of diurnal groundwater fluctuation on microbial degradation. Int J Phytoremediation. 11(5):509-523. https://doi.org/10.1080/15226510802656326

WHO. 2004. Toluene in drinking-water: background document for development of WHO *Guidelines for Drinking-water Quality*. World Health Organization. No. WHO/SDE/WSH/03.04/116.

Wilhelm R, Szeitz A, Klassen TL, Mohn WW. 2014. Sensitive, efficient quantitation of ^{13}C-enriched nucleic acids via ultrahigh-performance liquid chromatography-tandem mass spectrometry for applications in stable isotope probing. Appl Environ Microbiol. 80(23):7206-7211. https://doi.org/10.1128/AEM.02223-14

Wilhelm RC, Hanson BT, Chandra S, Madsen E. 2018. Community dynamics and functional characteristics of naphthalene-degrading populations in contaminated surface sediments and hypoxic/anoxic groundwater. Environ Microbiol. 20(10):3543-3559. https://doi.org/10.1111/1462-2920.14309

Williams LE, Cullen N, DeGiorgis JA, Martinez KJ, Mellone J, Oser M, Wang J, Zhang Y. 2019. Variation in genome content and predatory phenotypes between *Bdellovibrio* sp. NC01 isolated from soil and *B*. *bacteriovorus* type strain HD100. Microbiology. 165(12):1315-1330. https://doi.org/10.1099/mic.0.000861

Wilson J, Bartz R, Limmer M, Burken J. 2013. Plants as bio-indicators of subsurface conditions: impact of groundwater level on BTEX concentrations in trees. Int J Phytoremediation. 15(9):900-910. https://doi.org/10.1080/15226514.2013.765769

Wilson JL, Limmer MA, Samaranayake VA, Schumacher JG, Burken JG. 2017. Tree sampling as a method to assess vapor intrusion potential at a site characterized by VOC-contaminated groundwater and soil. Environ Sci Technol. 51(18):10369-10378. https://doi.org/10.1021/acs.est.7b02667

Winderl C, Anneser B, Griebler C, Meckenstock RU, Lueders T. 2008. Depth-resolved quantification of anaerobic toluene degraders and aquifer microbial community patterns in distinct redox zones of a tar oil contaminant plume. Appl Environ Microbiol. 74(3):792-801. https://doi.org/10.1128/AEM.01951-07

Winderl C, Schaefer S, Lueders T. 2007. Detection of anaerobic toluene and hydrocarbon degraders in contaminated aquifers using benzylsuccinate synthase (bssA) genes as a functional marker. Environ Microbiol. 9(4):1035-1046. https://doi.org/10.1111/j.1462-2920.2006.01230.x

Wolicka D, Suszek A, Borkowski A, Bielecka A. 2009. Application of aerobic microorganisms in bioremediation *in situ* of soil contaminated by petroleum products. Bioresour Technol. 100(13):3221-3227. https://doi.org/10.1016/j.biortech.2009.02.020

Wullschleger SD, Meinzer FC, Vertessy RA. 1998. A review of whole-plant water use studies in trees. Tree Physiol. 18(8-9):499-512. https://doi.org/10.1093/treephys/18.8-9.499

Xu H, Li X, Sun Y, Shi X, Wu J. 2016. Biodegradation of pyrene by free and immobilized cells of *Herbaspirillum chlorophenolicum* Strain FA1. Water, Air, & Soil Pollution. 227(4). https://doi.org/10.1007/s11270-016-2824-0

Yadav BK, Shrestha SR, Hassanizadeh SM. 2012. Biodegradation of toluene under seasonal and diurnal fluctuations of soil-water temperature. Water Air Soil Pollut. 223(7):3579-3588. https://doi.org/10.1007/s11270-011-1052-x

Yang S, Wen X, Shi Y, Liebner S, Jin H, Perfumo A. 2016. Hydrocarbon degraders establish at the costs of microbial richness, abundance and keystone taxa after crude oil contamination in permafrost environments. Sci Rep. 6:37473. https://doi.org/10.1038/srep37473

Yang S, Wen X, Zhao L, Shi Y, Jin H. 2014. Crude oil treatment leads to shift of bacterial communities in soils from the deep active layer and upper permafrost along the China-Russia Crude Oil Pipeline route. PLoS One. 9(5):e96552. https://doi.org/10.1371/journal.pone.0096552

Yao Y, Shen R, Pennel KG, Suuberg EM. 2013. A numerical investigation of oxygen concentration dependence on biodegradation rate laws in vapor intrusion. Environ Sci Process Impacts. 15(12):2345-2354. https://doi.org/10.1039/c3em00421j

Zalesny R, Wiese A, Bauer E, Riemenschneider D. 2006. Sapflow of hybrid poplar (*Populus nigra* L.×*P*. *maximowiczii* A. Henry 'NM6') during phytoremediation of landfill leachate. Biomass and Bioenergy. 30(8-9):784-793. https://doi.org/10.1016/j.biombioe.2005.08.006

Zhou J, Palumbo AV, Tiedje JM. 1997. Sensitive detection of a novel class of toluene-degrading denitrifiers, *Azoarcus tolulyticus*, with small-subunit rRNA primers and probes. Appl Environ Microbiol. 63(6):2384-2390.

Zwieniecki MA, Newton M. 1995. Roots growing in rock fissures - their morphological adaptation. Plant Soil. 172(2):181-187. https://doi.org/10.1007/Bf00011320

APPENDICES

Appendix A: Chapter 2 Supplementary Material

Toluene concentration and compound-specific carbon isotope analysis (CSIA)

Toluene concentration analysis in the soil vapor phase was performed using an Agilent 6890 gas chromatograph (GC) with DB5 column (60 m, 0.18 mm ID, 1 µm film thickness) coupled to an Agilent 5975 mass spectrometer (MS). Toluene concentration analysis in groundwater samples was conducted using an Agilent 7890A GC equipped with a DB-VRX column (20 m, 0.18 mm ID, 1 µm film thickness) and an electron capture detector (ECD). The carrier gas was helium with a flow rate of 1 mL/min. The GC oven temperature for toluene analysis was held at 36 °C for 4 min, then ramped at 16 °C/min to 85 °C, further ramped at 30 °C/min to 210 °C and held for 3 min. Compound-specific carbon isotope analysis (CSIA) of toluene in the vapor phase and in groundwater was conducted using a TRACE^{TM} GC and DeltaPlus XP isotope mass spectrometer (IRMS) via a combustion III interface (Thermo Finnigan, Germany). The accumulated toluene mass on the sorbent in WMS was first desorbed using an ATD 400 Thermal Desorber (Perking Elmer), whereby the desorption temperature was set at 250 °C for 1.5 min. The desorbed toluene mass was transferred to a cold trap set at -30 °C. The trap was heated to 280 °C to release the compounds to the GC column (DB-624, 60 m, 0.32 mm, 1.8 µm) for chromatographic separation. The toluene in groundwater samples was first pre-concentrated with a purge-and-trap system by purging 25 mL for ten minutes with N2 (40 mL/min) and trapping the toluene on a Vocarb 3000 trap. After desorption, toluene was transferred to the GC column (DB-624, 60 m, 0.32 mm, 1.8 µm) for chromatographic separation. The GC oven temperature for toluene analysis in the vapor phase and in groundwater was held at 40 °C for 2 min, then ramped at 15 °C/min to 200 °C and held for 2 min. The carrier gas was helium with constant flow of 2.4 mL/min. After separation the sample was conveyed to the IRMS for analyzing the carbon isotope signature relative

to a standard and the result was expressed in the delta notation (VPDB $\delta = R/R_{std} - 1)\cdot 1000$, where R and R_{std} are the isotope ratios of the sample and the standard, respectively.

Quantitative PCR

Total bacteria and toluene-degrading bacterial populations were assessed by quantifying genes and transcripts (cDNA) in metagenomic samples. All qPCR reactions were carried out on a CFX96™ Real-Time PCR Detection System (Bio-Rad Laboratories Inc.) and conformed to MIQE guidelines. Primers, target, and reaction conditions are detailed in Table S1. All DNA and cDNA samples were diluted to empirically-derived PCR inhibition-free dilutions prior to analysis according to previously-described spike inhibition testing methods (Reardon et al. 2013). Final reaction volumes of 20 µL included 10 µL of 2x SsoFast™ EvaGreen® Supermix, 0.5 µM of each primer, 2 µL (DNA) or 3 µL (cDNA) template, and nuclease- free sterile water. Standard curves for qPCR were produced by serially diluting plasmids containing target genes over a minimum dynamic range of 5 logs. Plasmid standards were constructed by cloning genes from environmental samples or pure culture DNA into TOPO TA plasmids (Life Technologies Corp.), and target gene identities were verified by sequencing Laboratory Services, University of Guelph, Guelph, Canada). Duplicate sample and duplicate/triplicate standard replicates were run for each target. A minimum of two no-template controls were run with each qPCR assay, yielding null or negligible values. Amplicon specificity was confirmed using melt curve analysis and/or by visualizing qPCR products on agarose gels. Single threshold Cq determination was used for qPCR data analysis. All assays performed had efficiencies between 95.5% and 103.7%, R^2 values between 0.996 and 1.000, and standard curve slopes between -3.434 and -3.237.

Table A.1: Degradation gene and transcript quantities and bacterial 16S rRNA gene quantities in rhizosphere soil (mean ± SE) and depth intervals of sampling trenches (m bgs).

Depth	ST1	ST2	[1] ST3	ST1	ST2	ST3
	PHE genes			PHE transcripts		
1	$2.29\times10^{6} \pm 4.18\times10^{5}$ [a]	$1.86\times10^{6} \pm 3.30\times10^{5}$	$3.82\times10^{6} \pm 1.10\times10^{6}$ [a]	$6.51\times10^{4} \pm 4.40\times10^{4}$	$1.02\times10^{5} \pm 7.56\times10^{4}$	$3.29\times10^{4} \pm 2.08\times10^{4}$
2	$1.41\times10^{6} \pm 1.67\times10^{5}$ [a,b]	$1.73\times10^{6} \pm 2.25\times10^{5}$	$1.05\times10^{6} \pm 1.42\times10^{5}$ [b]	$4.85\times10^{3} \pm 2.66\times10^{3}$ [‡]	$3.80\times10^{4} \pm 1.16\times10^{4}$ [†]	$5.68\times10^{3} \pm 1.13\times10^{3}$ [‡]
3	$6.32\times10^{5} \pm 5.18\times10^{4}$ [b,†]	$1.53\times10^{6} \pm 1.03\times10^{5}$ [‡]	$8.56\times10^{5} \pm 2.10\times10^{5}$ [b,†]	$1.56\times10^{4} \pm 9.38\times10^{3}$	$1.72\times10^{4} \pm 3.82\times10^{3}$	$2.53\times10^{4} \pm 1.22\times10^{4}$
	RMO genes			RMO transcripts		
1	$7.04\times10^{5} \pm 2.06\times10^{5}$	$5.31\times10^{5} \pm 1.04\times10^{5}$	$9.74\times10^{5} \pm 1.56\times10^{4}$	$2.79\times10^{4} \pm 7.77\times10^{3}$	$1.05\times10^{4} \pm 7.95\times10^{3}$	$8.25\times10^{3} \pm 4.13\times10^{3}$
2	$7.02\times10^{5} \pm 7.64\times10^{4}$	$8.65\times10^{5} \pm 9.70\times10^{4}$	$7.86\times10^{5} \pm 1.20\times10^{5}$	$2.45\times10^{4} \pm 1.61\times10^{4}$	$2.13\times10^{4} \pm 5.56\times10^{3}$	$3.18\times10^{3} \pm 3.18\times10^{3}$
3	$4.21\times10^{5} \pm 1.78\times10^{5}$	$6.45\times10^{5} \pm 8.56\times10^{4}$	$8.71\times10^{5} \pm 1.63\times10^{5}$	$2.69\times10^{4} \pm 1.35\times10^{4}$	$2.00\times10^{4} \pm 6.84\times10^{3}$	ND
	TOD genes			*bssA* genes		
1	$7.38\times10^{4} \pm 4.30\times10^{3}$ [†,‡]	$5.04\times10^{4} \pm 9.73\times10^{3}$ [†]	$1.08\times10^{5} \pm 6.63\times10^{3}$ [a,‡]	$6.75\times10^{6} \pm 1.41\times10^{6}$	$4.43\times10^{6} \pm 9.45\times10^{5}$	$8.64\times10^{6} \pm 3.84\times10^{5}$ [a]
2	$6.90\times10^{4} \pm 6.79\times10^{3}$	$7.58\times10^{4} \pm 6.81\times10^{3}$	$6.19\times10^{4} \pm 4.84\times10^{3}$ [b]	$6.03\times10^{6} \pm 3.56\times10^{5}$	$6.17\times10^{6} \pm 6.51\times10^{5}$	$4.84\times10^{6} \pm 3.50\times10^{5}$ [b]
3	$3.07\times10^{4} \pm 1.51\times10^{4}$	$4.90\times10^{4} \pm 7.36\times10^{3}$	$5.60\times10^{4} \pm 1.18\times10^{4}$ [b]	$2.58\times10^{6} \pm 1.07\times10^{6}$	$4.45\times10^{6} \pm 4.28\times10^{5}$	$4.59\times10^{6} \pm 8.15\times10^{5}$ [b]
	16S rRNA genes			Bottom of depth intervals (mbgs)		
1	$8.08\times10^{8} \pm 1.70\times10^{8}$ [a]	$4.67\times10^{8} \pm 1.14\times10^{8}$	$9.56\times10^{8} \pm 1.35\times10^{8}$ [a]	0.33	0.51	0.53
2	$5.29\times10^{8} \pm 3.30\times10^{7}$ [a,b]	$5.36\times10^{8} \pm 4.33\times10^{7}$	$3.70\times10^{8} \pm 4.26\times10^{7}$ [b]	0.97	1.14	1.22
3	$1.27\times10^{8} \pm 5.64\times10^{7}$ [b]	$2.74\times10^{8} \pm 9.41\times10^{6}$	$3.32\times10^{8} \pm 8.36\times10^{7}$ [b]	1.70	1.93	1.37

Values are mean ± standard error (n = 3) gene and transcript copies per gram dry rhizosphere soil of replicate samples. Values in the same column with dissimilar superscript lowercase letters and in the same row with dissimilar superscript symbols are significantly different ($P<0.05$) as assessed using the Holm-Šídák multiple comparisons test.

[1] One replicate excluded in depth 3 of ST3 for gene quantification (n = 2)

ND: not detected

Table A.2: Primers, targets, and reaction conditions used in quantitative real-time PCR reactions

Primer[a]	Sequence	Target	Expected product size (bp)	Cycle conditions	qPCR standard source[b]	Reference
16S 335r	5'-CADACTCCTACGGGAGGC	Bacterial 16S rRNA gene V3-V4 regions	429	Initial denaturation (98 °C, 2 min), 40 cycles of dissociation (98 °C, 7 s) and annealing/extension (59 °C, 15 s), and melt curve analysis (65 °C to 95 °C in 0.5 °C increments for 5 s)	Alphaproteobacteria environmental amplicon	(Dorn-In et al. 2015)
16S 769r	5'-ATCCTGTTTGMTMCCCVCRC					
RMO-F	5'-TCTCVAGCATYCAGACVGACG	Ring-hydroxylating monooxygenases	466	Initial denaturation (98 °C, 2 min), 40 cycles of dissociation (98 °C, 5 s) and annealing/extension (63 °C, 30 s), and melt curve analysis (65 °C to 95 °C in 0.5 °C increments for 5 s)	Toluene monooxygenase (*tmoA*-like) environmental amplicon	(Baldwin et al. 2003)
RMO-R	5'-TTKTCGATGATBACRTCCCA					
PHE-F	5'-GTGCTGACSAAYCTGYTGTTC	Phenol hydroxylases/ Ring-hydroxylating monooxygenases	203	Initial denaturation (98 °C, 2 min), 42 cycles of dissociation (98 °C, 10 s), annealing (49 °C, 7 s), and extension (60 °C, 3 s), and melt curve analysis (65 °C to 95 °C in 0.5 °C increments for 5 s)	Phenol hydroxylase environmental amplicon	(Baldwin et al. 2003)
PHE-R	5'-CGCCAGAACCAYTTRTC					
TOD-F	5'-ACCGATGARGAYCTGTACC	Toluene dioxygenases	754	Initial denaturation (98 °C, 2 min), 40 cycles of dissociation (98 °C, 10 s) and annealing/extension (60 °C, 30 s), and melt curve analysis (65 °C to 95 °C in 0.5 °C increments for 5 s)	*Pseudomonas putida* F1 (DSM 6899) amplicon	(Baldwin et al. 2003)
TOD-R	5'-CTTCGGTCMAGTAGCTGGTG					

bssA 7772f	5’-GACATGACCGACGCSATYCT	Benzyl succinate synthase alpha subunit	800	Initial denaturation (98 °C, 2 min), 40 cycles of dissociation (98 °C, 7 s) and annealing/extension (63 °C, 35 s), and melt curve analysis (65 °C to 95 °C in 0.5 °C increments for 5 s)	*Thauera aromatica* (DSM 6984) amplicon	(von Netzer et al. 2013)
bssA 8543r	5’-TCGTCRTTGCCCCAYTTNGG					
M13-F	5’-TTGTAAAACGACGGCCAG					
M13-R	5’-GGAAACAGCTATGACCAT					

[a] M13 primers were used in spike PCR inhibition testing for empirical derivation of inhibition-free sample dilutions (Reardon et al. 2013).

[b] Bacterial 16S rRNA, RMO, and PHE genes were cloned from soil samples obtained from the research site; all standard gene identities were verified by sequencing.

Table A.3: Degradation gene and transcript quantities and bacterial 16S rRNA gene quantities in rhizosphere soil (mean ± SE) and depth intervals of sampling trenches (m bgs).

	Rhizoplane soil			Roots		
Depth	[1]ST1	ST2	ST3	ST1	ST2	ST3
	RMO genes			RMO genes		
1	$2.68\times10^{6} \pm 2.96\times10^{5}$ [a]	$1.83\times10^{6} \pm 5.62\times10^{5}$	$3.32\times10^{6} \pm 5.61\times10^{5}$	$1.96\times10^{7} \pm 5.90\times10^{6}$	$1.45\times10^{7} \pm 3.08\times10^{6}$	$8.07\times10^{6} \pm 1.27\times10^{6}$
2	$2.50\times10^{6} \pm 1.01\times10^{5}$ [a]	$3.03\times10^{6} \pm 1.75\times10^{5}$	$2.75\times10^{6} \pm 2.89\times10^{4}$	$5.81\times10^{6} \pm 1.07\times10^{6}$	$2.24\times10^{7} \pm 1.09\times10^{7}$	$2.08\times10^{7} \pm 8.69\times10^{6}$
3	$1.09\times10^{6} \pm 4.25\times10^{5}$ [b,†]	$3.23\times10^{6} \pm 2.12\times10^{5}$ [‡]	$1.94\times10^{6} \pm 1.23\times10^{5}$ [†]	$1.55\times10^{7} \pm 4.77\times10^{6}$	$2.19\times10^{7} \pm 5.01\times10^{6}$	$1.49\times10^{7} \pm 1.07\times10^{6}$
	TOD genes			TOD genes		
1	$1.22\times10^{5} \pm 7.00\times10^{3}$ [a]	$1.08\times10^{5} \pm 1.81\times10^{4}$	$1.29\times10^{5} \pm 2.40\times10^{4}$	$8.14\times10^{5} \pm 1.57\times10^{5}$ [†]	$6.27\times10^{5} \pm 4.88\times10^{4}$ [†,‡]	$3.35\times10^{5} \pm 4.04\times10^{4}$ [b,‡]
2	$1.32\times10^{5} \pm 9.57\times10^{3}$ [a]	$1.48\times10^{5} \pm 7.32\times10^{3}$	$1.23\times10^{5} \pm 8.60\times10^{3}$	$4.44\times10^{5} \pm 7.54\times10^{4}$	$1.04\times10^{6} \pm 3.96\times10^{5}$	$1.14\times10^{6} \pm 1.28\times10^{5}$ [a]
3	$8.46\times10^{4} \pm 2.91\times10^{3}$ [b]	$1.31\times10^{5} \pm 1.37\times10^{4}$	$1.11\times10^{5} \pm 1.27\times10^{4}$	$4.47\times10^{5} \pm 7.99\times10^{4}$	$6.61\times10^{5} \pm 1.89\times10^{5}$	$9.09\times10^{5} \pm 2.00\times10^{5}$ [a,b]
	PHE genes			[2]PHE genes		
1	$4.19\times10^{6} \pm 3.72\times10^{5}$ [a]	$3.11\times10^{6} \pm 9.99\times10^{5}$	$4.37\times10^{6} \pm 6.61\times10^{5}$ [a]	$1.87\times10^{7} \pm 4.77\times10^{6}$	NQ	$5.75\times10^{6} \pm 7.94\times10^{5}$
2	$2.97\times10^{6} \pm 1.74\times10^{5}$ [b]	$5.90\times10^{6} \pm 2.77\times10^{5}$	$3.17\times10^{6} \pm 4.16\times10^{5}$ [a,b]	5.34×10^{6}	NQ	1.42×10^{7}
3	$1.39\times10^{6} \pm 4.95\times10^{5}$ [c,†]	$8.66\times10^{6} \pm 2.07\times10^{6}$ [‡]	$1.98\times10^{6} \pm 3.74\times10^{5}$ [b,†]	$1.06\times10^{7} \pm 3.06\times10^{6}$	$3.88\times10^{7} \pm 2.16\times10^{7}$	$7.85\times10^{6} \pm 1.95\times10^{6}$
	16S rRNA genes			16S rRNA genes		
1	$9.38\times10^{8} \pm 9.16\times10^{7}$ [a]	$8.81\times10^{8} \pm 3.10\times10^{8}$	$1.02\times10^{9} \pm 2.15\times10^{8}$	$4.44\times10^{9} \pm 1.31\times10^{9}$	$3.81\times10^{9} \pm 7.12\times10^{8}$	$1.61\times10^{9} \pm 8.15\times10^{7}$
2	$8.16\times10^{8} \pm 3.08\times10^{7}$ [a]	$1.74\times10^{9} \pm 3.57\times10^{8}$	$8.49\times10^{8} \pm 2.30\times10^{8}$	$1.94\times10^{9} \pm 1.75\times10^{8}$	$5.64\times10^{9} \pm 3.63\times10^{9}$	$4.23\times10^{9} \pm 7.58\times10^{8}$
3	$2.30\times10^{8} \pm 2.97\times10^{7}$ [b]	$1.32\times10^{9} \pm 9.03\times10^{7}$	$6.39\times10^{8} \pm 9.92\times10^{7}$	$3.52\times10^{9} \pm 6.83\times10^{8}$	$4.21\times10^{9} \pm 1.42\times10^{9}$	$2.96\times10^{9} \pm 6.30\times10^{8}$

Values are mean ± standard error (n = 3) gene copies per gram pelleted rhizoplane soil or gram dry root of replicate samples. Values in the same column with dissimilar superscript lowercase letters and in the same row with dissimilar superscript symbols are significantly different ($P<0.05$) as assessed using the Holm-Šídák multiple comparisons test.

[1] One replicate excluded in depth 3 of ST1 for rhizoplane gene analysis (n = 2)

[2] ST1 and ST3 endosphere samples at depth 2 had two replicates in which PHE genes were present but not quantifiable due to non-specific PCR product (n = 1)

NQ: Root samples in which PHE genes were present but not quantifiable due to non-specific PCR product in all replicates

Appendix B: Chapter 3 Supplementary Material

Figure B.1: Phytoremediation stand. Planting *ca*. 2008 (left), 220 m^2 fenced area indicated by red dashed box (center), stand *ca*. 2016 (right).

		
May 20, 2017	May 28, 2017	June 10, 2017
		
June 20, 2017	June 30, 2017	July 11, 2017
		
July 30, 2017	August 21, 2017	September 14, 2017

		
October 3, 2017	October 31, 2017	

Figure B.2: Tree canopy conditions during transpiration study period. Photos collected using a Moultrie Game Camera.

Amplicon library preparation and sequencing

Amplicon library preparation for 16S *rRNA* genes in root tissue DNA samples was performed as described previously (Habtewold et al. 2018). Amplicons were prepared targeting the V4 region of the 16S *rRNA* gene and performed using a semi-nested approach in two PCR steps: (1) initial amplification and addition of Illumina adapter sequences, and (2) addition of Illumina index tags to the amplicons from step (1). In the first PCR step, 16S *rRNA* genes were amplified using the modified 515F (Parada) and 806R (Apprill) primer sets (Table S1), as described in the Earth Microbiome Project 16S Illumina Amplicon Protocol (Caporaso et al. 2012; Apprill et al. 2015; Parada et al. 2016), with Illumina adapter sequences (see Supplemental Material for primer sequences). Reactions were performed in duplicate using a Thermo Scientific™ Phusion™ Hot Start II High-Fidelity DNA Polymerase kit (ThermoFisher Scientific Cat. No. F-549L) in 25 µL reactions containing: 5 µL 5X Phusion™ HF Buffer, 0.75 µL 100% DMSO, 0.25 µL Phusion™ Hot Start II DNA Polymerase, 1.25 µL of each primer (10 µM), 0.5 µL Thermo Scientific™ dNTP 10 mM Mix (ThermoFisher Scientific Cat. No. R0192), and 2 µL of 20x diluted template. This PCR inhibition-free sample dilution was empirically-derived using plasmid spike testing as described by Reardon et al. (2013). Reaction conditions were: initial denaturation (98 °C, 3 min), 25 cycles of dissociation (98 °C, 10 s), annealing (50 °C, 30 s), and extension (72 °C, 15 s), and final extension (72 °C, 10 min). Duplicate reactions for each sample were pooled and products were purified using Wizard® 511

SV Gel and PCR Clean-Up System (Promega) silica spin columns. Illumina index tags then were added to the ends of amplicons from step (1) using an Illumina Nextera® 515 XT DNA Library Preparation Kit (Illumina Inc.), with different combinations of forward and reverse index primers for each sample. Reaction conditions and reagent proportions were as in the first PCR step, except the total reaction volume was 50 µL, a template volume of 4 µL purified amplicons was used, and 8 PCR cycles were performed. PCR products were purified using an Agencourt AmPure XP magnetic bead purification kit (Beckman Coulter) with a resuspension volume of 25 µL and tested for correct amplicon length using gel electrophoresis. Multiplexed sample sequencing was performed at the University of Guelph Advanced Analysis Centre, Genomics Facility (Guelph, ON) using a MiSeq Reagent Kit v2 (500-cycle) (Illumina Inc.), producing paired end reads of 250 bp in length. Prior to sequencing, libraries were normalized using a SequalPrep™ Normalization Plate Kit (ThermoFisher Scientific) and library quality was assessed using a Bioanalyzer DNA1000 chip (Agilent).

Table B.1: Primer sequences and cycle conditions for step 1 of 16S *rRNA* amplicon preparation. Template-complementing sequences in bold, adapter sequences underlined.

Primer	Sequence	Cycle conditions
515F (Parada) with Adapter A:	5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG**GTGYCAGCMGCCGCGGTAA**-3'	initial denaturation (98 °C, 3 min), 25 cycles of dissociation (98 °C, 10 s), annealing (50 °C, 30 s), and extension (72 °C, 15 s), and final extension (72 °C, 10 min)
806R (Apprill) with Adapter B:	5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG**GGACTACNVGGGTWTCTAAT**-3'	

Table B.2: Sequencing quality control metrics for DADA2 pipeline

Sample	Replicate	input	filtered	denoised	merged	non-chimeric
ST1	1	194262	134392	134392	115364	108622
	2	183922	108800	108800	91941	83079
ST2	1	168875	120971	120971	109043	95010
	2	173309	118448	118448	103981	89909
ST3	1	239288	150755	150755	135185	127907
	2	216744	146295	146295	132281	117863
	3	232922	168724	168724	151057	135957

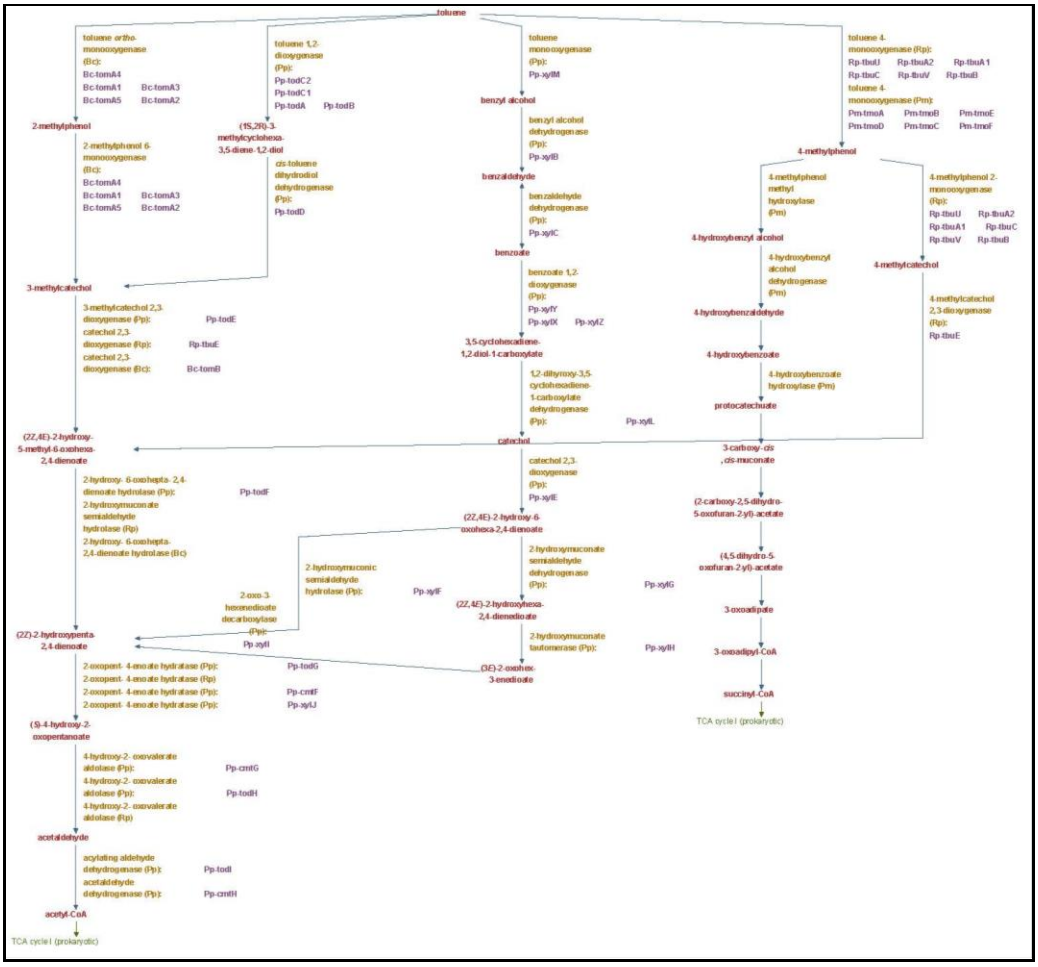


Figure B.3: MetaCyc superpathway for aerobic toluene degradation (PWY-5183) with compounds (red), enzyme names (yellow), and gene names (purple). Two-letter prefixes on purple-colored gene names indicate host organisms.

Table B.3: *In planta* toluene concentrations. ND = non-detect, NQ = non-quantifiable.

Tree	Toluene (µg/L)	Diameter 0.5 m ags
A2	0.26	26.7
A5	1.2	21.8
A9	0.11	20.2
B7	0.63	17.4
C4	ND	18
C9	ND	16.9
D1	ND	16.4
D5	NQ	15.7
E4	ND	16.7
F2	NQ	22.3
G1	ND	19.6
G3	0.19	18.5
G6	0.79	18.1
Control 1	ND	20.05
Control 2	ND	19.1

Table B.4: Genera with significantly greater bacterial 16S *rRNA* sequence abundance in high toluene (ST3) compared to low toluene (ST1/2) roots. Differences in relative abundances assessed using Welch's t-test.

Phylum	Class	Order	Family	Genus	Relative abundance ± stdev (%)		Difference between means (%)	p-values
					High	Low		
Actinobacteria	Actinobacteria	Streptomycetales	*Streptomycetaceae*	*Streptomyces*	9.632 ± 1.056	6.641 ± 1.515	2.991	0.048
Proteobacteria	Alphaproteobacteria	Rhizobiales	*Xanthobacteraceae*	*Ancylobacter*	0.329 ± 0.076	0.016 ± 0.018	0.313	0.025
Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	*Rubritaleaceae*	*Luteolibacter*	0.487 ± 0.117	0.193 ± 0.081	0.294	0.048
Actinobacteria	Actinobacteria	Micromonosporales	*Micromonosporaceae*	*Catelliglobosispora*	0.109 ± 0.015	0.06 ± 0.01	0.049	0.023
Proteobacteria	Deltaproteobacteria	Myxococcales	*Polyangiaceae*	*Pajaroellobacter*	0.044 ± 0.005	0.009 ± 0.009	0.036	0.002
Bacteroidetes	Bacteroidia	Chitinophagales	*Chitinophagaceae*	*Flavihumibacter*	0.047 ± 0.008	0.015 ± 0.017	0.031	0.044

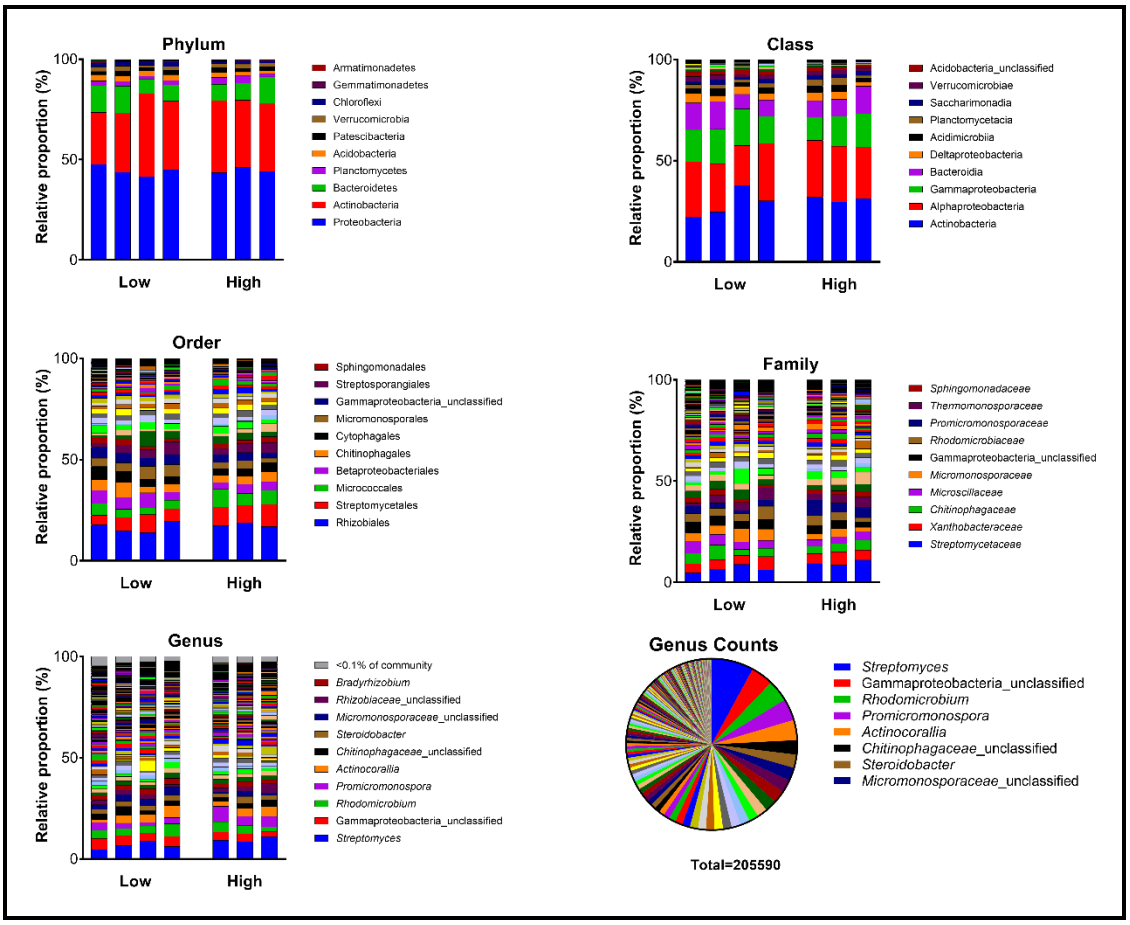
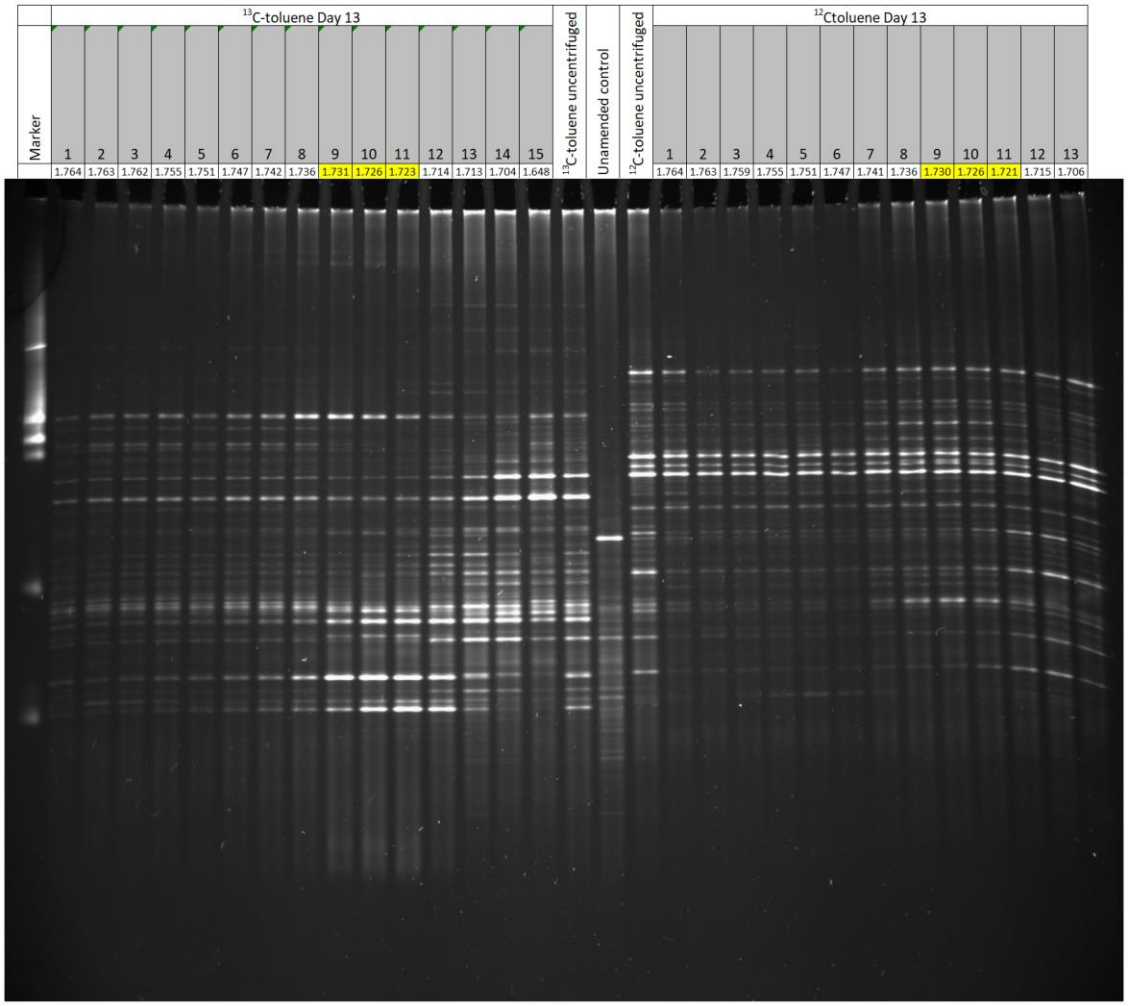


Figure B.4: Relative bacterial 16S *rRNA* sequence abundances over phylum, class, order, family, and genus taxonomic levels in high toluene (ST3) and low toluene (ST1/2) roots, and total genus-level sequences counts for all samples. Legends indicate major taxa at each level. Sequences with <0.1% relative abundance on genus level grouped together.

Appendix C: Chapter 5 Supplementary Material

A)



B)

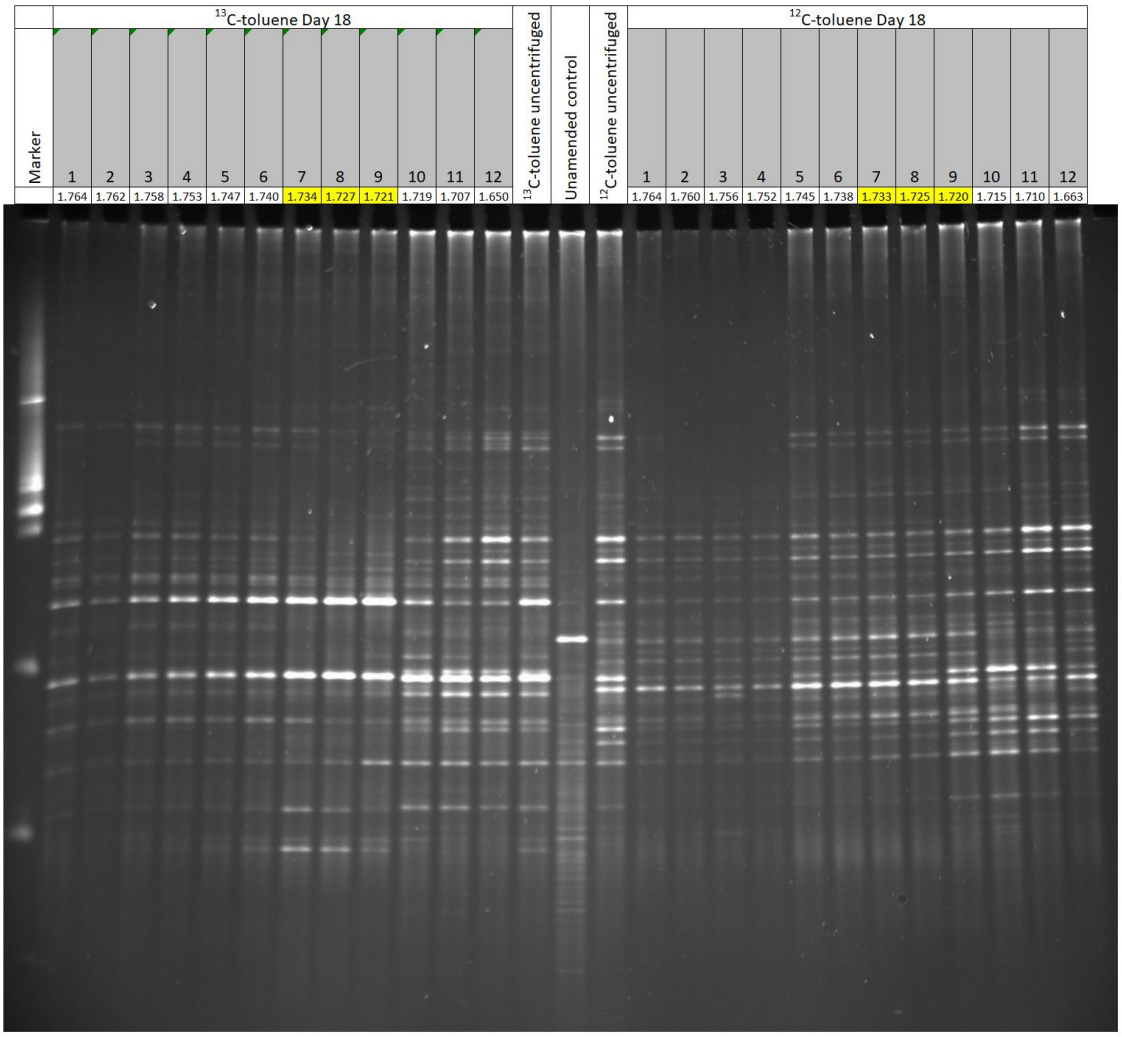


Figure C.1: DGGE profiles of bacterial/archaeal 16S *rRNA* genes amplified from density-resolved fractions of DNA extracted from toluene-amended microcosms on day 13 (A) and day 18 (B) of toluene incubation. Densities of fractions indicated below fraction numbers, with labeled buoyant density fractions highlighted in yellow. DGGE profiles for non-density-resolved toluene amended and unamended DNA samples are shown as well. A 100 bp DNA marker (Promega Corp. G210A) was used as reference marker in the first lane of each gel.